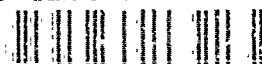AD-A238 576

2514628-PH

②

Adaptive Liquid Crystal TV Based Joint Transform Correlator
as Applied to Real-Time Pattern Recognition

Final Report

By:

F.T.S. Yu
Principal Investigator

Department of Electrical and Computer Engineering
The Pennsylvania State University
University Park, PA 16802

Prepared for:

Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

Attention: Bobby D. Guenther

Contract DAAL03-87-K-0147
ARO Proposal Number 25146-PH

Period Covered: October 1, 1987 to March 31, 1991

Date: May 23, 1991

91-05552

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services. Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204 Arlington, VA 22202-4302 and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 23, 1991	3. REPORT TYPE AND DATES COVERED Final Report - 10/1/87 - 3/31/91

4. TITLE AND SUBTITLE
Adaptive Liquid Crystal TV Based Joint Transform
Correlator as Applied to Real-Time Pattern Recognition

5. FUNDING NUMBERS

6. AUTHOR(S)
Francis T.S. Yu

DAAL03-87-K-0147

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Electrical and Computer Engineering
The Pennsylvania State University
University Park, PA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER

NA

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARO 25146.28-PH

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The primary goal of this research is to study a programmable joint-transform correlator (JTC) using liquid crystal television (LCTV) panels for adaptive real-time pattern recognition applications. The technique can improve the pattern recognition and identification technology that is of interest to the U.S. Army. The technique we studied is a real-time programmable electro-optical architecture. There are several reasons for selecting the optical technique over their digital and electronic counterparts, as follows: Optical technique is capable of handling a large space-bandwidth image; optical technique is capable of performing parallel operations; optics can perform massive interconnections; optical transformation can be operated at high speed, etc. By using the LCTV, the pattern under observation can be correlated with a large number of recallable image memories. In addition, the LCTV technique is rather simple and economical to operate. The LCTV-optical correlator, in principle, can be designed into a compact portable form for insitu application. Brief outlines of the major research findings and publications are provided in this report.

14. SUBJECT TERMS Joint transform correlator, liquid crystal television correlator, optical pattern recognition			15. NUMBER OF PAGES 164
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

NSN 7540-01 280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

FINAL REPORT

1. ARO PROPOSAL NUMBER: 25146-PH
2. PERIOD COVERED BY REPORT: July 1, 1988 - March 31, 1991
3. TITLE OF PROPOSAL: Adaptive Liquid Crystal TV Based Joint
                             Transform Correlator as Applied to Real-Time
                             Pattern Recognition
4. CONTRACT OR GRANT NUMBER: DAAL03-87-K-0147
5. NAME OF INSTITUTION: The Pennsylvania State University
6 AUTHORS OF REPORT: Francis T.S. Yu

CONTENTS   Page

CONTENTS (Cont.)   Page

7. LIST OF MANUSCRIPTS SUBMITTED OR PUBLISHED UNDER ARO SPONSORSHIP DURING
    THIS REPORTING PERIOD, INCLUDING JOURNAL REFERENCES:

F.T.S. Yu and T. Nagata, "Binary Phase only Joint Transform Correlator,"
Microwave and Optical Technology Letters, Vol. 2, pp. 15-19, January
1989.

F. T. S. Yu, S. Jutamulia, R. V. Yelamarty and D. Gregory, "Adaptive
Joint Transform Correlator for Real-Time Color Pattern Recognition,"
Optics and Laser Technology, Vol. 21, pp. 189-192, June 1989.

F.T.S. Yu, C. Zhang, Y. Jin and S. Jutamulia, "Application of
Holographic Associative Memory to a Hybrid Binary Adder," SPIE
Proceeding on Holographic Optics: Design and Application, Vol. 883, pp.
254-259, January 1988.

F.T.S. Yu, S. Jutamulia and E. Tam, "Gray Level Pseudocolor Encoding
using a Liquid Crystal Television," Journal of Optics, Vol. 19, pp.
129-123, May-June 1989.

F.T.S. Yu, F. Cheng, T. Nagata, and Don A. Gregory, "Effects of Fringe
Binarization on Multi-Object Joint Transform Correlation," Applied
Optics, Vol. 28, pp. 2988-2990, August 1989.

D. A. Gregory, J. A. Loudin and F.T.S. Yu, "Illumination Dependence of
the Joint Transform Correlation," Applied Optics, Vol. 28, pp.
3288-3290, August 1989.

F.T.S. Yu, Q. W. Song, Y. S. Cheng and D. A. Gregory, "Comparison of
Detection Efficiencies for Vander Lugt and Joint Transform Correlators,"
Applied Optics, Vol. 29, pp. 225-232, January 1990.

F.T.S. Yu, X. Li, E. Tam, S. Jutamulia and D. A. Gregory, "Rotational
Invariant Pattern Recognition with a Programmable Joint Transform
Correlator," Proceeding of 1988 Conference on Pattern Recognition for
Advance Missile Systems, November 14-15, 1988.

T. W. Lu, S. D. Wu, X. Xu, and F.T.S. Yu, "A 2-D Programmable Optical
Neural Network," Applied Optics, Vol. 28, pp. 4908-4913, November 1989.

F.T.S. Yu, X. Y. Li, E. Tam, S. Jutamulia and D. A. Gregory, "Rotation-
Invariant Pattern Recognition with a Programmable Joint Transform
Correlator," Applied Optics, Vol. 28, pp. 4725-4727, November 1989.

X. J. Lu, F.T.S. Yu and D. A. Gregory, "Analytical Comparison of Vander
Lugt and Joint Transform Optical Correlators," Applied Physics B, Vol.
51, pp. 153-164, August 1990.

F.T.S. Yu, E. Tam, and D. A. Gregory, "High Efficient Joint Transform
Correlator," Optics Letters, Vol. 15, pp. 1029-1031, September 1990.

F.T.S. Yu, C. H. Zhang, Y. Jin and D. A. Gregory, "Nonconventional Joint Transform Correlator," Optics Letters, Vol. 14, pp. 922-924, September 1989.

F.T.S. Yu, Y. Jin and C. Zhang, "Symbolic Logic Processing using Cascaded LCTVs," Microwave and Optical Technology Letters, pp. 309-313, September 1989.

T. Lu, S. Wu, X. Xu and F.T.S. Yu, "Optical Implementation of Programmable Neural Network," Proceeding of SPIE on Optical Pattern Recognition, Vol. 1053, January 1989.

F.T.S. Yu, E. Tam, X. Li, and D. A. Gregory, "Detection of Rotational and Scale Varying Objects with a Programmable Joint Transform Correlator," Proceeding of SPIE on Optical Pattern Recognition, Vol. 1053, January 1989.

E. C. Tam, F.T.S. Yu and D. A. Gregory, "Target Tracking with an Adaptive Joint Transform Correlator," Proceedings on SPIE Conference on Aerospace Pattern Recognition, Vol. 1098, March 1989.

S. Wu, T. W. Lu, X. Xu and F.T.S. Yu, "An Adaptive Optical Neural Network using a High Resolution Video Monitor," Microwave and Optical Technology Letters, Vol. 2, pp. 252-257, July 1989.

T. W. Lu, X. Xu, S. Wu and F.T.S. Yu, "A Neural Network Model using Inter-Pattern Association (IPA)," Applied Optics, Vol. 29, pp. 284-288, January 1990.

T. W. Lu, X. Xu, S. Wu and F.T.S. Yu, "A Neural Network Model using Inter-Pattern Association (IPA)," Proceeding of the International Joint Conference on Neural Networks, Vol. I, p. 596, June 18-22, 1989.

F. Cheng, X. Xu, S. Wu, F.T.S. Yu and D. A. Gregory, "Restoration of Linear-Motion-Blurred Images using a Joint Transform Processor," Proceeding of the SPIE Conference on Optical Information Processing, Aug. 9-10, 1989.

X. Y. Li, F.T.S. Yu and D. A. Gregory, "Optical Implementation of the Kittler Transform for Image Classification," Proceeding of the SPIE Conference on Optical Information Processing, Aug. 9-10, 1989.

S. Wu, Q. Song, A. Mayers, F.T.S. Yu and D. A. Gregory, "Reconfigurable Interconnections Using Nonlinear Photorefractive Crystal", Applied Optics, Vol. 29, pp. 1118-1125, March 1990.

F. Cheng, X. Yu, S. Wu, F.T.S. Yu and D. A. Gregory, "Restoration of Images Blurred Due to Linear Motion Using a Joint Transform Procession," Microwave and Optical Technology Letters, Vol. 3, pp. 21-27, January 1990.

E. Tam, S. Wu, A. Tanone, F.T.S. Yu and D. A. Gregory, "Closed-Loop Binary Phase Connection of a LCTV Using a Point Diffraction Interferometer," IEEE Photonics Technology Letters, Vol. 2, pp. 143-146, February 1990.

T. Lu, X. Xu, S. Wu and F.T.S. Yu, "Hetero Association Neural Network for Pattern Recognition," Proceedings of 1989 IEEE International Conference on Systems, Man, and Cybernetics, November 14-17, 1989.

E. C. Tam, F.T.S. Yu, D. A. Gregory, and R. D. Juday, "Autonomous Real-Time Object Tracking with an Adaptive Joint Transform Correlator," Optical Engineering, Vol. 29, pp. 314-320, April 1990.

F.T.S. Yu, T. W. Lu, X. Y. Yang, and D. A. Gregory, "Compact Size Neural Network using Liquid Crystal Televisions " Proceeding of the SPIE Conference on Optical Computing, Vol. 1215, January 1990.

F.T.S. Yu, T. W. Lu, X. Y. Yang, and D. A. Gregory, "Optical Neural Network with Pocket-Size Liquid-Crystal Televisions," Optics Letters, Vol. 15, pp. 863-865, August 1990.

D. A. Gregory, J. A. Loudin, J. C. Kirsch, E. C. Tam, and F.T.S. Yu, "Application of Hybrid Modulating Properties of the Liquid Crystal Television," Applied Optics (In Press).

T. W. Lu, F.T.S. Yu and D. A. Gregory, "Self-Organizing Neural Network for Unsupervised Learning," Optical Engineering, Vol. 29, pp. 1107-1113, September 1990.

E. C. Tam, F.T.S. Yu, A. Tanone, D. A. Gregory, and R. Juday, "Data Association Multiple Target Tracking using a Phase-Mostly LCTV," Optical Engineering, Vol. 29, pp. 1114-1121, September 1990.

F.T.S. Yu, E. C. Tam, D. A. Gregory, and R. Juday, "Multi-target Tracking with a Hybrid Joint Transform Correlator," Proceeding of the SPIE Conference on Real-Time Image Processing II, Vol. 1295, pp. 128-137, April 1990.

T. Lu, F.T.S. Yu, and D. A. Gregory, "Optical Neural Network for Unsupervised Learning," Proceeding of the SPIE Conference on Advances in Optical Information Processing IV, Vol. 1296, April 1990.

F.T.S. Yu, T. Lu and X. Yang, "Optical Implementation of Hetero-Association Neural Network with Inter-Pattern Association Model," International Journal of Optical Computing (In Press).

F.T.S. Yu, X. Yang and T. Lu, "Space-Time Sharing Optical Neural Network," Optics Letters, Vol. 16, pp. 247-249, February (1991).

4

J. C. Kirsch, D. A. Gregory, J. A. Loudin, E. C. Tam and F.T.S. Yu, "Light Efficient Joint Transform Optical Correlator," Proceedings of the SPIE Conference on Optical Information Processing Systems and Architectures II, San Diego, CA, July 8-13, 1990.

F.T.S. Yu, T. Lu and D. A. Gregory, "Compact Optical Neural Network using Liquid Crystal Television," Proceeding of the ICO-15th Conference on Optics in Complex Systems, Garmisch, Germany, SPIE Vol. 1319, pp. 200-201, August 1990.

F.T.S. Yu and E. C. Tam, "A High Efficiency Joint Transform Correlator for Symbolic Substitution," Proceeding of the ICO-15 Conference on Optics in Complex Systems, Garmisch, Germany, August 5-10, 1990

F.T.S. Yu, "Optical Neural Network: Architecture, Design and Models," Invited Proceeding of IEEE TENCON '90 Conference on Computer and Communication Systems, Hong Kong, pp. 12-16, September 24-27, 1990.

F.T.S. Yu and T. Lu, "Adaptive Optical System for Neural Computing," Proceeding of IEEE TENCON '90 Conference on Computer and Communication Systems, Hong Kong, September 24-27, 1990.

F.T.S. Yu, S. Wu, A. Mayers, S. Rajan and D. A. Gregory, "Color Holographic Storage in $LiNbO_3$," Optics Communications, Vol. 81, pp. 343-352, March 1991.

F.T.S. Yu, S. Wu, A. Mayers, S. Rajan, "Wavelength Multiplexed Reflection Matched Spatial Filters using $LiNbO_3$," to Optics Communications, Vol. 81, pp. 343-347, March 1991.

F.T.S. Yu, T. W. Lu and D. A. Gregory, "Self-Learning Optical Neural Network," Proceedings of Spatial Light Modulators and Applications, Technical Digest, Vol. 14, pp. 24-27, September 1990.

F.T.S. Yu, S. Wu, A. Mayers, S. Rajan and D. A. Gregory, "Color Holographic Storage in $LiNbO_3$," Proceedings of Spatial Light Modulators and Applications, Technical Digest, Vol. 14, pp. 153-156, September 1990.

X. Yang, T. Lu, F.T.S. Yu and D. A. Gregory, "Redundant Interconnection Interpattern Association Neural Network," Applied Optics (In Press).

F.T.S. Yu, Y. Li, X. Yang, T. Lu and D. A. Gregory, "Application of Moment Invariant Pattern Recognition to Optical Neural Net," submitted to Optik.

8. SCIENTIFIC PERSONNEL SUPPORTED BY THIS PROJECT AND DEGREES AWARDED DURING THIS REPORTING PERIOD:

    8.1 Scientific Personnel

        F.T.S. Yu   - Principal Investigator
        A. Mayers   - Research Assistant
        E. Tam   - Research Assistant
        X. Li   - Research Assistant
        F. Cheng   - Research Assistant
        A. Tanone   - Research Assistant
        S. D. Wu   - Research Assistant
        S. Rajan   - Research Assistant
        T. W. Lu   - Research Assistant

    8.2 Degrees Awarded

        "Real-Time Optical Vander Lugt and Joint Transform Correlation Systems," Xiao-Jing Lu, Ph.D. Degree, completed August 1988.

        "Applications of Moment Invariants to Neurocomputing for Pattern Recognition," Yajun Li, Ph.D., completed June 1990.

        "Modulation properties of a twisted nematic liquid crystal spatial light modulator and its applications in a joint transform correlator," E. C. Tam, Ph.D., completed August 1990.

        "Study of an optical neural network...architecture, design and models.," T. W. Lu, Ph.D., completed August 1990.

        "Optical image classification using optical/digital hybrid image processing system," J. X. Li, completed October 1990.

9. BRIEF OUTLINE OF RESEARCH FINDINGS

    In a conventional coherent optical signal processor [1], the processing operation is usually carried out at the spatial frequency or Fourier plane with a complex spatial filter [2]. This type of coherent optical processor offers a myriad of complicated processing operations [3]. Its success is primarily due to the profound diffraction phenomena. However, complex signal processing can also be achieved by the spatial impulse response using a joint transform processor. There are several inherent advantages of using the joint Fourier transform processor as compared with the conventional coherent

processor: (1) spatial filter adjustment is not imposed; (2) a higher input space-bandwidth product; (3) generally, a higher modulation index of the joint transform hologram; (4) lower spatial carrier frequency, etc. In view of these advantages, a joint transform processor, in principle, is capable of performing optical signal processing more efficiently, particularly in the application of real-time pattern recognition. The purpose of this research program is to investigate an adaptive liquid crystal TV based correlator as applied to real-time pattern recognition and tracking.

In this period, from July 1, 1988 to March 31, 1991, we have accomplished several major tasks on the research of real-time pattern recognition with an LCTV based correlator for which various results have been reported in the refereed journals and conference proceedings; with the approval of the Department of the U.S. Army. Sample copies of these publications are included in this final technique report in the subsequent sections, to provide a concise documentation of our findings. In the following sections, we shall give an overview of our research work done during this program. We will highlight some of the accomplished works. A list of publications resulting from this support has been cited in the preceding section.

9.1   Rotation Invariant Pattern Recognition (APPENDIX 10.1)

During this research program, we have successfully investigated rotational invariant pattern recognition using a programmable liquid crystal television (LCTV) based joint transform correlator (JTC). We have introduced two methods to recognize targets with different in-plane rotational orientations. One method is using real value implementation of circular harmonic expansion techniques for joint transform target detection, the other

is using an information reduction principle for multi-reference image generation such that rotational objects can be detected. Exceedent experimental demonstrations for these two methods have been performed. A paper of this nature is published in Applied Optics [4].

9.2   Effects on Fringe Binarization (APPENDIX 10.2)

We have also studied the effects of fringe binarization on multi-object joint transform correlation. We have shown that binarization joint transform interference fringes for multiple targets would produce false alarms and misses. We have concluded that fringe binarization may not be usefully applied for multi-target recognition, although it can be applied successfully for noisefree singer target correlation. A letter to the editor of these findings was published in Applied Optics [5].

9.3   Detection Efficiencies (APPENDIX 10.3)

In this period of research, we investigated the detection efficiencies for Vander Lugt and joint transform correlators. We have found that the relative performance of a Vander Lugt correlator would generally perform better in the multi-object and under noisy environment. These effects are caused by the signal dependent nature of the joint transform filter. Unlike the Vander Lugt filter, it is signal independent. However, JTC has the advantage in real-time implementation, alleviates the spatial filter synthesis and alignment problems. We have published a paper of these findings in Applied Optics [6].

9.4   JTC Color Pattern Recognition (APPENDIX 10.4)

In this research program, we have also developed an adaptive joint transform correlator for real-time pattern recognition using a color LCTV. We have shown that this color JTC is capable of exploiting the spatial and the

spectral content of the target. A paper of this nature is published in the Optics and Laser Technology [7].

9.5   Comparison of Vander Lugt and Joint Transform Correlator (APPENDIX 10.5)

We completed a theoretical study of comparing the Vander Lugt and joint transform correlators. Multiplicative noise and differences between input and reference functions are considered so that their effects on optical correlation can be estimated. We have found, in general, the JTCs are easier to construct, simpler to handle and control and more practical for application than Vander Lugt correlators. However, the basic distinction between these two correlators are: the JTC filter is signal dependent, whereas the Vander Lugt filter is signal independent. A paper of this study is published in Applied Physics [8].

9.6   LCTV Color Encoder (APPENDIX 10.6)

While studying the polarization properties of the liquid crystal materials, we have developed a technique to perform real-time gray level pseudocolor encoding using a LCTV. The LCTV is used under two different polarizer/analyzer combination to generate a positive and a negative image. These images are then encoded with two primary colors by positioning the color filter respectively. The superposition of these images would produce a gray level pseudocolor encoded image. The result of this study has been published in Journal of Optics [9].

9.7   White-Light JTC (APPENDIX 10.7)

In this period of research, we have also developed a technique of performing joint transform correlation using a white-light source. The temporal and spatial coherence requirement of a joint transform correlator (JTC) have been studied. Excellent experimental results have been obtained.

Since the white-light source emanates all visible spectra, the white-light JTC is capable of exploiting the spectral content of the target. A paper of this nature has been published in Optics Letters [10].

S.8   High-Efficient JTC (APPENDIX 10.8)

We have developed a high efficient JTC. By modifying the joint transform power spectrum of the input objects, using a spectral sampling method, the read out light, as well as the physical area, of the square low detector can be fully utilized. As a result, the output correlation intensity can be substantially increased. This technique would benefit the real-time optical pattern recognition performance. To improve the signal-to-noise, the read-out can also be done by using partially coherent light   A paper of this report is published in Optics Letters [11].

9.9   Autonomous Target Tracking (APPENDIX 10.S)

One of the most interesting research tasks we have accomplished in this period is the application of liquid crystal TV band JTC to autonomous target tracking. The fundamental element of the system is a real-time optical joint transform correlator using a liquid-crystal television. The parallelism, high processing speed, and adaptive properties of this optical system assure high correlation between objects in two sequential frames. The relative position of the object can then be determined based on the location of the correlation peak. System performance is elevated and experimental demonstrations have been conducted. We stress that this technique has important applications to real-time missile, aircraft and vehicle tracking, which are very significant for the U.S. Army strategic defense needs in guided weapons and high speed target tracking systems. We have reported a paper of this research in Optical Engineering [12].

9.10. Image Deconvolution (APPENDIX 10.10)

In this phase of research, we have also performed image deconvolution due to linear motion using a 1-D joint transform correlator. The technique is basically using a blur function, as the reference object, in a joint transform correlator to reverse the negative phase spectrum of the blurred image. In other words, the method is equivalent to using a deblurring phase filter in a conventional coherent processor. The advantages of this technique are simplicity and flexibility, for which blurred images can be corrected by using real-time devices. A paper reports the image deblurring using a 1-D JTC architecture was published in Optics Communications [13], and a paper deals with the restoration of linear smeared image is published in Microwave and Optical Technology Letters [14].

9.11 Reconfigurable Interconnections (APPENDIX 10.11)

High speed computing has aroused great interest in optical interconnections. We have, in this period, investigated a reconfigurable optical interconnection using a nonlinear photorefractive crystal. We have used the coupled wave theory and law of refraction, to analyze the volume of holograms. Reconfigurable interconnections are discussed employing either wavelength tuning or spatial division technique. Reflection type volume holograms can be used for a large number of reconfigurable interconnections in terms of finite wavelength tunability. Transmission volume holograms encoded in pinhole holograms can be easily reconfigurated by SLM for optical interconnections. Experimental demonstrations have also been conducted in this phase of research. We stress that, the proposed reconfigurable interconnection technique, would have significant impact on the future research of high speed optical computing. We have reported a paper to the

Applied Optics of this study [15].

9.12   Binary Phase Correction (APPENDIX 10.12)

In this research program, binary phase correlation of a liquid crystal television (LCTV) using a point diffraction interferometer has been investigated. The use of a point diffraction interferometer offers the simplicity and advantages of a common-path interferometer for examining the phase non-uniformity of the LCTV, while the on-line and closed-loop architecture is capable of generating the optimum phase correction function. Both computer simulations and experimental results demonstrate the effectiveness of the binary phase correction scheme. The results obtained from the bipolar phase only joint transform correlation reveals the feasibility of this phase correction technique in actual applications. Mention must be made that this binary phase correction scheme can be used to compensate for the phase distortion of a general optical signal processing system. Other phase modulating real-time addressable SLMs which have a higher contrast ratio and a larger space-bandwidth product, such as the magnetooptic SLMS or liquid crystal light valves may be employed to produce even better results. A paper reporting these findings is published in the IEEE Photonics Technology Letters [16].

9.13   Image Classification (APPENDIX 10.13)

We have used a joint transform correlator for the study of image classification by the Kittler-Young transform. The Kittler-Young transform, is a nonparametric method for feature extraction, which can be effectively applied to image classification. The advantage of using the K-Y transform is that it goes one step further in utilizing the eigenvalues. The eigenvalues are separated into two parts, which are proportional to the class means and

variances respectively. The effects on the feature's discriminative power are in opposing ways. To overcome the difficulty caused by this contradiction, a feature selection criterion based solely on the class means after normalizing class variances is used with the K-Y transform. Some of the limitations of using the first three algorithms can be alleviated using the K-Y transform, resulting in a more discriminative feature space. Experiments to classify birds and fish, using the K-Y transform are demonstrated. The results show that the K-Y transform offers a high feature selection than using the Fukunaga-Koontz transform. We have published a paper of these findings in the Micro. Opt. Tech. Lett. [17].

9.14   IPA Neural Network (APPENDIX 10.14)

During this research program, we investigated a neural network model using Inter-Pattern Association (IPA). The IPA model uses basic logical operations to determine the inter-pattern association (i.e., association between the reference patterns), and simple logical rules are applied to construct tri-state interconnections in the network. Computer simulations for the reconstruction of similar English letters embedded in the random noise by IPA model have shown improved performance in comparison with the Hopfield model. A 2-D hybrid optical neural network is used to demonstrate the usefulness of the IPA model. Since there are only three gray levels used in the Interconnection Weight Matrix (IWM) for IPA model, the dynamic range imposed on a Spatial Light Modulator (SLM) is rather relaxed, and the interconnections are much simpler than the Hopfield model. A paper of this study is published in Appl. Optics. [18].

9.15   LCTV Neural Network (APPENDIX 10.15)

We have developed an 8x8 neuron optical neural network using pocket size

liquid crystal televisions (LCTVs). The size of the optical neural network is about 75x15x25 cm^3. The operation speed of this LCTV neural network, with 8x8 neurons is about 1.2×10^5 interconnection operations/sec. However, for a 16x20 neural work, the operation speed can be as high as 3×10^5 inter.oper./sec. Experimental results from this LCTV neural network are demonstrated. A paper of this work is published in Optics Letters [19].

9.16 Nonconventional JTC (APPENDIX 10.16)

In this research program, we have also investigated a nonconventional joint transform correlator (NJTC). The advantages of the NJTC are the efficient use of the light source, the use of smaller transform lenses, higher correlation peaks, and a higher carrier fringe frequency. Since NJTC can, in principle, process all the information that a conventional optical processor can offer, we believe that the NJTC would have all the processing capabilities of a coherent processor. A paper of this study is published in Optics Letters [20].

9.17 Multi-Target Tracking (Appendix 10.17)

One of the major advantages of optical signal processing is the parallelism in handling high space-bandwidth signals. However, the price paid is the complexity and stringent alignment of an optical system. On the other hand, its digital counterpart offers flexibility and programmability while sacrificing full parallelism. We have investigated a technique of using data association target tracking in a motion sequence via an adaptive joint transform correlator. The massive data in the field of view can be reduced to a few correlation peaks. The average velocity of a target during the tracking cycle is then determined from the location of the correlation peak. We have

used a data association algorithm for the analysis of these correlation

signals, for which multiple targets can be tracked. Simplicity in optical

architecture, relaxation of alignment requirements, and the adaptive property

are the major advantages of a hybrid JTC in this mode of operation. The phase

modulation property of a LCTV has been investigated and was employed in this

experiment to improve the detection efficiency of a JTC. A motion sequence of

three targets was used as the input scenes to demonstrate the applicability of

this technique and satisfactory results were obtained. A paper reporting the

performance of this technique is published in Optical Engineering [21].

9.18 Unsupervised Learning Neural Net (APPENDIX 10.18)

One of the features in neural computing must be the adaptability to

changeable environment and to recognize unknown objects. We have, in this

period of research, implemented Kohonon's self-organizing feature map for

unsupervised learning in an optical neural network. A compact optical neural

network of 64 neurons using liquid crystal televisions is used for this study.

To test the performances of the self-organizing neural network, experimental

demonstrations with computer simulations are conducted. Effects due to

unsupervised learning parameters have been analyzed. It is shown that the

optical neural network is capable of performing both unsupervised learning and

pattern recognition operations simultaneously, by setting two matching scores

in the learning algorithm. We have also implemented the fast and slower

learning models in the neural networks. We have found that by using a slower

learning rate the construction of the memory matrix becomes topologically more

organized. Moreover, by introducing the forbidden regions in the memory

space, it would enable the neural network to learn new patterns without

erasing the old ones. A paper describing these findings is published in

Optical Engineering [22].

9.19 Color Holography Using LiNbO₃ (APPENDIX 10.19)

   Two of the most widely used white-light holograms must be the reflection hologram of Denisyuk and the rainbow hologram of Benton. In reflection hologram, a thickness emulsion of about 20μm would have a wavelength selectivity about $\Delta\lambda/\lambda$ = 1/40, which is high enough to produce color hologram images without significant color blur. However, the physical requirements for constructing a reflection hologram is rather stringent, which prevents its wide spread use of applications. On the other hand, construction of a rainbow hologram requires a narrow slit, for which the parallax information of the hologram image would be partly lost. In this research program, we have also demonstrated that color holograms can be constructed in a photorefractive crystal using a thick LiNbO₃ crystal with a "white-light" laser. Since photorefractive crystal is much thicker than conventional photographic emulsion, it provides a higher wavelength selectivity such that the color blur can be minimized. Furthermore, the construction of photorefractive holograms is in real-time mode and the shrinkage of the emulsion can be prevented. As in contrast with the photographic film, multiplexing color holograms in a photorefractive crystal is feasible for which we have experimentally demonstrated. We have shown that to minimize the color crosstalk, a high wavelength selectivity (i.e., thick crystal) photorefrac media than normally required should be used. A paper reporting the color holographic imaging using a white-light laser is published in Optics Communications [23].

9.20 Compact LCTV Neural Network (APPENDIX 10.20)

   In this research, we have also developed a compact optical neural network using two tightly cascaded liquid crystal televisions (LCTVs). This

new optical architecture offers the compactness in size, easy alignment, higher light efficiency, better image quality, and low cost. The implementation of the auto-associative and hetero-associative memories to this new architecture are also tested. The most import .. `spect of this new architecture is that a large numerical aperture : .. `(' is used in the previous optical neural network can be avoided. 'l.t . a more compact size system can be built. Moreover, the lightly cascade `/‾Vs of the system also relax the fine alignment problem posed by the opt. . system. This proposed architecture with experimental demonstration is published in Applied Optics [24].

9.21 Wavelength Multiplexed Matched Filters (APPENDIX 10.21)

We have studied the wavelength multiplexing capability using a thick LiNbO$_3$ for reflection matched spatial filter synthesis. In principle, over 1000 spatial filters can be multiplexed in a 4mm thick photorefractive crystal. The wavelength multiplexing technique in a crystal filter has the advantage that neither the object nor the reference beam needs to be changed in the writing or reading processes, even though the two processes can be used in conjunction to further increase the multiplexing capabilities. Although this technique has been used in holographic film which matched filters were multiplexed, the thin film emulsion leads to low wavelength selectivity, and the filters are difficult to produce. By referring the coupled-mode analysis, we have analyzed the wavelength multiplexing capability of reflection type matched spatial filters. By integrating over the thickness of the crystal, we have shown that thick crystal filters exhibit high wavelength selectivity as well as retaining a good degree of the shift invariance. Thus, by recording single wavelength filters, and reading-out the filters with a multi-wavelength

source, parallel pattern recognition can be performed. A paper on this study is published in Optics Communications [25].

9.22 Space-Time-Sharing Neural Network (APPENDIX 10.22)

To provide the neural network with learning ability, spatial light modulators (SLMs), such as liquid crystal television (LCTV), have been used as the programmable devices. In a fully interconnect neural network, every neuron at the input plane is connected to all the neurons at the output plane. For example, 1000 neurons would require a million interconnections. Thus, a very high resolution SLM is required for the massive interconnection. However, the resolution of the currently available LCTVs is rather limited, which poses an obstacle for large scale operation we have, in this phase of research, used a space-time sharing technique to alleviate this constraint. Since the LCTV neural network is essentially a white-light processing system, by partitioning the interconnection weight matrix into an array of submatrices, a large space-bandwidth pa er (SBP) can be processed with a smaller neural network. We have shown that to achieve a large SBP of the system, additional expenditure of processing time is needed. The amount of processing time increases as the square function of the SBP of the input pattern. We have also experimentally demonstrated that a 12x12 and 24x24 element pattern can be indeed on a 6x6 LCTV neuron network. A paper of this study is published in Optics Letters [26].

9.23 Binary Phase Only JTC (APPENDIX 10.23)

We have also in the research investigate a joint transform hybrid optical correlator as applied to binary phase only filtering. The interference fringes of the joint Fourier transformation, due to the reference and the input targets, are captured by a CCD camera feeding into a

microcomputer for binarization. A binarized phase only filter, in principle, can be written onto a spatial light modulator (SLM). By simple coherent illumination of SLM, correlation operation of the reference and the input targets can be obtained computer simulated results for target embedded in an additive noise is provided. We have shown that high correlation peaks can be readily achieved under noisy environment. Since JTC filter is signal dependent, for multi-target recognition, it would produce false alarms and misses. A paper of this nature is published in Micro. Opt. Tech. Lett. [27].

9.24 Remarks

The ability to process a large quantity of information at a high speed makes the optical correlator an attractive candidate for applications to machine vision, target tracking and detection, etc. Although conceptually simple, the Vander Lugt correlator, which employs a holographic spatial filtering technique, has inherent filter synthesis and alignment problems, that prevent its wide spread practical application. On the other hand, the Joint Transform Correlator (JTC) is a simple and practical processor that overcomes these two major disadvantages. In this report, we have performed the major tasks proposed on the story of the LCTV joint transform correlator and other LCTV based system, as applied for pattern recognition. We are confident that our goal of realizing a practical adaptive LCTV based JTC for real-time recognition and identification would happen in the near future.

References

1. L. J. Cutrona, E. N. Leith, C. J. Palermo, and L. J. Porcello, "Optical Data Processing and Filtering Systems," IRE Trans. Inform. Theory, IT-6, 386 (1960).

2. A. Vander Lugt, "Signal Detection by Complex Spatial Filtering," IEEE Trans. Inform. Theory, IT-10, 139 (1964).

3. F.T.S. Yu, Optical Information Processing, Wiley-Interscience, New York, 1983.

4. F.T.S. Yu, X. Li, E. Tam, S. Jutamulia and D. A. Gregory, "Rotational Invariant Pattern Recognition with a Programmable Joint Transform Correlator," Appl. Opt. 28, 4725 (1989).

5. F.T.S. Yu, F. Cheng, T. Nagata, and D. A. Gregory, "Effects of Fringe Binarization of Multiobject Joint Transform Correlation," Appl. Opt., 28, 2988 (1989).

6. F.T.S. Yu, Q. W. Song, Y. S. Cheng and D. A. Gregory, "Comparison of Detection Efficiencies for Vander Lugt and Joint Transform Correlators," Appl. Opt., 29, 225 (1990).

7. F.T.S. Yu, S. Jutamulia, R. V. Yelamarty and D. A. Gregory," Adaptive Joint Transform Correlator for Real-Time Color Pattern Recognition," Optics and Laser Technology, 21, 189 (1989).

8. X. J. Lu, F.T.S. Yu, and D. A. Gregory, "Comparison of Vander Lugt and Joint Transform Correlators," Appl. Phy. B51, 153 (1990).

9. F.T.S. Yu, S. Jutamulia and E. Tam, "Gray Level Pseudocolor Encoding Using a Liquid Crystal Television," J. Opt., 19, 129 (1988).

10. F.T.S. Yu and Y. S. Cheng, "White-Light Joint-Transform Correlator," Opt. Lett., 15, 192 (1990).

11. F.T.S. Yu, E. C. Tam and D. A. Gregory, "High-Efficiency Joint-Transform Correlator," Opt. Lett., 15, 1029 (1990).

12. E. C. Tam, F.T.S. Yu, D. A. Gregory and R. D. Juday, "Autonomous Real-Time Object Tracking with an Adaptive Joint-Transform Correlator," Opt. Eng. 29, 314 (1990).

13. F.T.S. Yu and Y. S. Chen, "Deblurring Using Joint Transform Processor," Opt. Commun., 74, 19 (1989).

14. F. Cheng, X. Xu, S. Wu, F.T.S. Yu, and D. A. Gregory, "Restoration of Blurred Images Due to Linear Motion Using a Joint Transform Processor," Micro. Opt. Tech. Lett., 3, 24 (1990).

15. S. Wu, Q. Song, A. Mayers, D. A. Gregory and F.T.S. Yu, "Reconfigurable Interconnections Using Photorefractive Holograms," Appl. Opt., 29, 1118 (1990).

16. E. C. Tam, S. Wu, A. Tanone, F.T.S. Yu and D. A. Gregory, "Closed-Loop Binary Phase Correction of an LCTV Using a Point Diffraction Interferometer," IEEE Phot. Tech. Lett., 2, 143 (1990).

17. J. X. Li, F.T.S. Yu and D. A. Gregory, "Image Classification by the Kittler-Yo Transform Using a Joint-Transform Correlator," Micro. Opt. Tech. Lett., 4, 207 (1991).

18. T. Lu X. Xu, S. Wu, and F.T.S. Yu, "Neural Network Model Using Interpattern Association," Appl. Opt. 29, 284 (1990).

19. F.T.S. Yu, T. Lu, X. Yang and D. A. Gregory, "Optical Neural Network with Pocket-Sized Liquid Crystal Televisions," Opt. Lett., 15, 863 (1990).

20. F.T.S. Yu and C. Zhuang, Y. Jin and D. Gregory, "Nonconventional Joint-Transform Correlator," Opt. Lett., 14, 922 (1989).

21. E. C. Tam, F.T.S. Yu, A. Tanone, D. A. Gregory and R. D. Juday, "Data Association Multiple Target Tracking Using a Phase-Mostly Liquid Crystal Television," Opt. Eng., 29, 1114 (1990).

22. T. Lu, F.T.S. Yu and D. A. Gregory, "Self-Organizing Optical Neural Network for Unsupervised Learning," Opt. Eng., 29, 1107 (1990).

23. F.T.S. Yu, S. Wu, A. Mayers, S. Rajan and D. A. Gregory, "Color Holographic Storage in LiNbO$_3$," Opt. Commun., 81, 348 (1991).

24. X. Yang, T. Lu and F.T.S. Yu, "Compact Optical Neural Network Using Cascaded Liquid Crystal Television," Appl. Opt., 29, 5223 (1990).

25. F.T.S. Yu, S. Wu, A. Mayers, S. Rajan, "Wavelength Multiplexed Reflection Matched Filters Using LiNbO$_3$," Opt. Commun., 81, 343 (1991).

26. F.T.S. Yu, X. Yang and T. Lu, "Space-Time-Sharing Optical Neural Network," Opt. Lett., 16, 247 (1991).

27. F.T.S. Yu and T. Nagata, "Binary Phase Only Joint Transform Correlator," Micro. Opt. Tech. Lett., 2, 15 (1989).

APPENDIX 10.1

Rotation Invariant Pattern Recognition

Rotation invariant pattern recognition with a programmable joint transform correlator

Francis T. S. Yu, Xiaoyang Li, Eddy Tam, Suganda Jutamulia, and Don A. Gregory

Don Gregory is with U.S. Army Missile Command, Research, Development, & Engineering Center, Research Directorate, Redstone Arsenal, Alabama 35898-5248; Suganda Jutamulia is with Quantex Corporation, 2 Research Court, Rockville, Maryland 20850; the other authors are with Pennsylvania State University, Department of Electrical Engineering, University Park, Pennsylvania 16802. Received 7 October 1988.

A circular harmonic expansion filter is real value implemented in a joint transform correlator architecture to perform rotation invariant pattern recognition.

One of the most common problems in pattern recognition is to identify objects of different sizes, angular orientations and perspectives. Utilizing the architecture of a joint transform correlator 'JTC'[1-3] as shown in Fig. 1, we propose to use the real value implementation of circular harmonic functions as references in a JTC to perform rotation invariant pattern recognition.

If a 2-D function $f(r,\theta)$ is continuous and integrable over the region $(0,2\pi)$, it can be expanded into a Fourier series as

$$f(r,\theta) = \sum_{m=-\infty}^{+\infty} F_m(r) \exp(im\theta), \qquad (1)$$

where

$$F_m(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r,\theta) \exp(-im\theta)d\theta. \qquad (2)$$

$F_m(r,\theta) = F_m(r) \exp(im\theta)$ is called the mth-order circular harmonic.

Previous research[4-7] has shown that by using a single harmonic $F_k(r,\alpha)$ as a complex reference function, the intensity of the correlation center is independent of the rotation angle α. Although $f(r,\theta)$ is a complex function in general, it is usually represented by a real quantity in practice. Thus Eq. (1) can be written as

$$f(r,\theta) = F_0(r) + 2 \sum_{m=1}^{+\infty} \mathrm{Re}[F_m(r) \exp(im\theta)]$$

$$= F_0(r) + 2 \sum_{m=1}^{+\infty} |F_m(r)| \cos(m\theta + \phi_m), \qquad (3)$$

where $F_m(r) = |F_m(r)| \exp(i\phi_m)$. In this real value implementation technique, we define the following two reference functions:

$$R_{k1}(r,\theta) = 2|F_k(r)| \cos(k\theta + \phi_k), \qquad (4)$$

$$R_{k2}(r,\theta) = 2|F_k(r)| \sin(k\theta + \phi_k). \qquad (5)$$

These real value functions are basically the real and imaginary parts of the circular harmonic $F_k(r,\theta)$. Using R_{k1} as the reference image function and $f(r,\theta + \alpha)$ as the input target function, the center value of the cross correlation becomes

$$C_{k1} = \int_0^{2\pi} \int_0^{\infty} f(r,\theta + \alpha) R_{k1} r\, dr\, d\theta. \qquad (6)$$

By substituting Eqs. (4) and (5) into Eq. (6), we have

$$C_{k1}(\alpha) = \int_0^{2\pi} \int_0^{\infty} \left\{ F_0(r) + 2 \sum_{m=1}^{\infty} |F_m(r)| \right.$$

$$\times \cos[m(\theta + \alpha) + \phi_m] \Big\}$$

$$\times \{2|F_k(r)| \cos(k\theta + \phi_k)\} r\, dr\, d\theta. \qquad (7)$$

Since all the terms in Eq. (7) vanish except the $m = k$ term, the above equation can be reduced to

$$C_{k1}(\alpha) = 4\pi \cos(k\alpha) \int_0^{\infty} r|F_k(r)|^2 dr$$

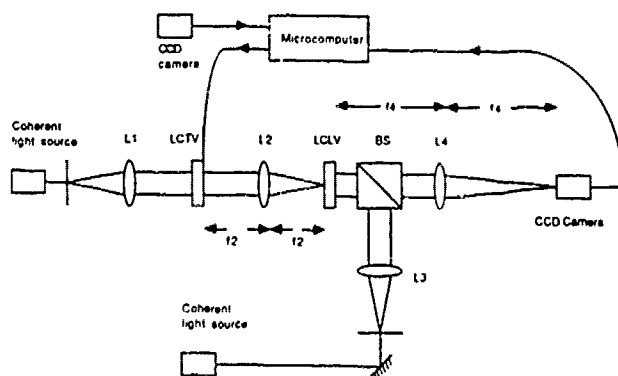$$= 2C_k(0) \cos(k\alpha). \qquad (8)$$

Fig. 1. Real-time programmable joint transform correlator: $L1$ and $L3$ are collimating lenses, $L2$ and $L4$ are Fourier transform lenses, and BS is a beam splitter.

Fig. 2. Input image used in the correlation experiment. The real and imaginary circular harmonic reference images are displayed at the upper part, and the input target is at the lower part.
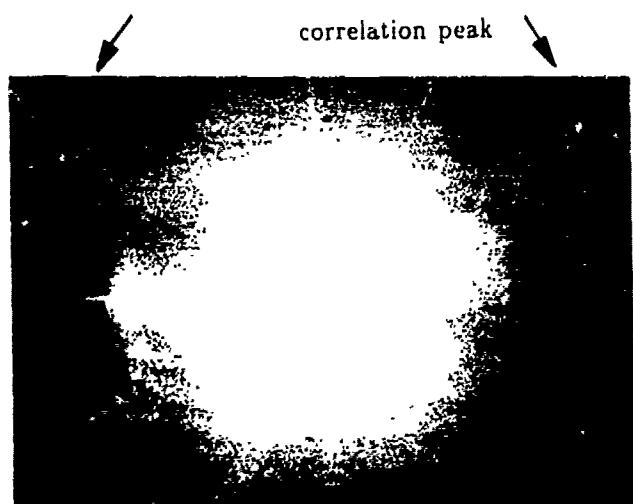

correlation peak

Fig. 3. Experimental result obtained from the JTC.

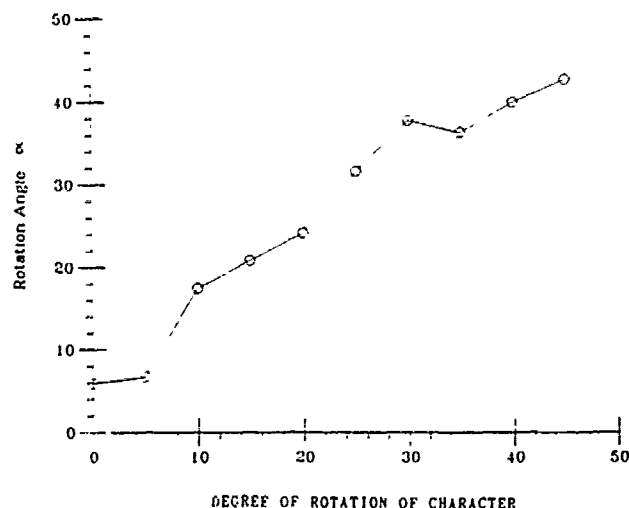DEGREE OF ROTATION OF CHARACTER
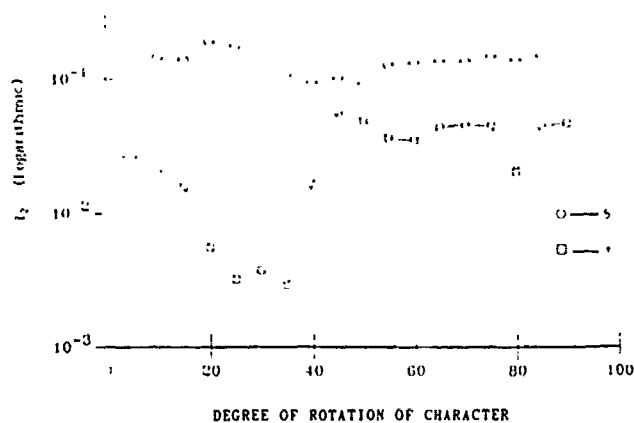

DEGREE OF ROTATION OF CHARACTER

Fig. 4. (a) Plot of the resultant peak intensity vs the input target orientation for objects 5 (upper curve) and 3 (lower curve), respectively. (b) Plot of the calculated rotation angle α vs input target orientation.

The corresponding intensity is

$$I_{k1}(\alpha) = 4C_k^2(0)\cos^2(k\alpha),\qquad(9)$$

which varies as a cosine square with respect to the target orientations. Likewise, if $R_{k2}(r,\theta)$ is used as the reference image function, the intensity of the correlation center is

$$I_{k2}(\alpha) = 4C_k^2(0)\sin^2(k\alpha).\qquad(10)$$

Thus, by adding the correlation intensities given in Eqs. (9) and (10), we obtain a new value

$$I_k = I_{k1}(\alpha) + I_{k2}(\alpha) = 4C_k^2(0),\qquad(11)$$

which is independent of the orientation angle α. Therefore, we conclude that rotation invariant correlation is achieved. Using Eqs. (9) and (10), the value of α can be determined as

$$\alpha = \frac{1}{k}\tan^{-1}\sqrt{I_{k1}/I_{k2}}.\qquad(12)$$

However, I_{k1} and I_{k2} are positive defined quantities; hence the value of α determined Eq. (12) is confined to region $[0,|\pi/$

$(2k)]$. Consequently, for each α thus computed, there arc $4k$ possible orientations of the target.

To insure a non-negative value, bias terms can be added to R_{k1} and R_{k2}. The positive definite images can then be written onto the input SLM, in which the undesired bias can be removed at the frequency plane. On the other hand, it is very difficult to implement a complex function onto a SLM. However, in a recent paper by Leclerc et al.,[7] it is shown that by using a binary phase-only technique, a circular harmonic expansion filter can also be implemented directly on a SLM in a VanderLugt correlator architecture.

Figure 2 consists of the target image 5 located in the lower half, and the real and imaginary parts of the second-order CHE of the character are located in the upper half. The input pattern recorded on a photographic film was then inserted at the input plane of the JTC shown in Fig. 1, and the two correlation peak intensities with respect to the real and imaginary CHE were recorded at the output plane. Figure 3 shows the output correlation result when the target is at 0°. Figure 4(a) shows the sum of these two intensities as a function of the target rotation. We can see that the peak

intensity is independent of the angular rotation of the target. Therefore, rotation invariant pattern recognition is achieved. When another numerical character 3, which has great similarity to the character 5, was used as the target image, the correlation response shows a high degree of fluctuation. This illustrates the discrimination of this filter. Furthermore, the orientation α of the target was also calculated by using Eq. (12). The plot of α vs the target orientation presented in Fig. 4(b) shows that the experimental result obtained is within an accuracy of $\pm 5°$.

It must be mentioned that this technique requires summation of two correlation peak intensities for each input target. This is a major drawback of this technique and introduces certain degrees of difficulties in practice. One of the two peaks might fall below the noise level of the detected signal and would be impossible to detect. However, the complementary peak in this situation emerges to give a high correlation peak intensity based on the equation that $\sin^2\theta + \cos^2\theta = 1$. Therefore, one of the two peaks would be adequate to detect the target in this situation.

When two or more targets are present at the input scene, multiple correlation signals are produced. Therefore, one also has to ensure that proper pairs of correlation peaks are to be added. We note that the space invariant property of a JTC provides one with the position information of the target based on the location of the correlation peak. Therefore, if a peak is located at position A on the output plane, its complimentary peak must be located at a location dictated by the value of A and the predefined relative position of the two references. The correct complementary peak's intensity can then be obtained accordingly.

In summary: The real value implementation of a circular harmonic function in a JTC has been investigated. By using the real and imaginary parts of circular harmonic reference functions, rotation and shift invariant object detection can be carried out with a JTC. The advantage of this technique is that the reference images are a positive real value, which can easily be implemented on programmable SLMs.

References

1. C. S. Weaver and J. W. Goodman, "A Technique for Optically Convolving Two Functions," Appl. Opt. 5, 1248–1250 (1966).
2. F. T. S. Yu and X. J. Lu, "A Real-Time Programmable Joint Transform Correlator," Opt. Commun. 52 (10 Nov. 1984)
3. F. T. S. Yu, S. Jutamulia, T. W. Lin, and D. A. Gregory. "Adaptive Real-Time Pattern Recognition Using a Liquid Crystal TV Based Joint Transform Correlator," Appl. Opt. 26, 1370–1372 (1987).
4. Y.-N. Hsu and H. H. Arsenault, "Optical Pattern Recognition Using Circular Harmonic Expression," Appl. Opt. 21, 4016–4019 (1982).
5. H. H. Arsenault, C. Ferreira, M. P. Levesque, and T. Szpolik, "Simple Filter with Limited Rotation Invariance," Appl. Opt. 25, 3230–3234 (1986).
6. Y.-N. Hsu, H. H. Arsenault, and G. April, "Rotation-Invariant Digital Pattern Recognition Using Circular Harmonic Expansion," Appl. Opt. 22, 4012–4015 (1982).
7. L. Leclerc, Y. Sheng, and H. H. Arsenault, "Rotation Invariant Phase-Only and Binary Phase-Only Correlation," Appl. Opt. 28, 1251–1256 (1989).

APPENDIX 10.2

Effects on Fringe Binarization

Effects of fringe binarization of multiobject joint transform correlation

Francis T. S. Yu, Feng Cheng, Toshio Nagata, and Don A. Gregory

Don Gregory is with U.S. Army Missile Command, Redstone Arsenal, Alabama 35898; Toshio Nagata is with Japan Defense Agency, Technical Research & Development Institute, Fifth Research Center, 3-13-1 Nagase, Yokosuka 239, Japan; and the other authors are with Pennsylvania State University, Electrical Engineering Department, University Park, Pennsylvania 16802.
Received 28 November 1988.

A study of the effect of fringe binarization on the joint transform correlator has shown that the interference fringes for multiple targets could produce false alarms and misses.

One of the important features of an optical correlator is based on the shift invariant property of the Fourier transformation which allows multiple targets to be detected and tracked. In an optical joint transform correlator (JTC),[1-3] both the reference object and the input targets are located in the input plane of the coherent optical processor. The corresponding joint transform power spectrum of the input objects (i.e., the reference object and input targets) consists of various carrier interference fringes, which can be extracted by a square law device such as an image detector. By inverse Fourier transforming these interference fringes, the cross correlations between reference object and input targets are produced.

In this Letter, the effects of binarization of the power spectrum[4] in a JTC are discussed. We assume that the reference object $f_0(x,y)$ and the input targets $f_i(x,y)$, $i = 1, 2,\dots,N$ are positioned at (a_i,b_i), $i = 0,1,\dots,N$ as shown in Fig. 1. With coherent illumination, the complex light distribution at the Fourier plane P_1 is given by

$$G(p,q) = \sum_{i=0}^{N} F_i(p,q) \exp[-j(a_i p + b_i q)], \qquad (1)$$

where

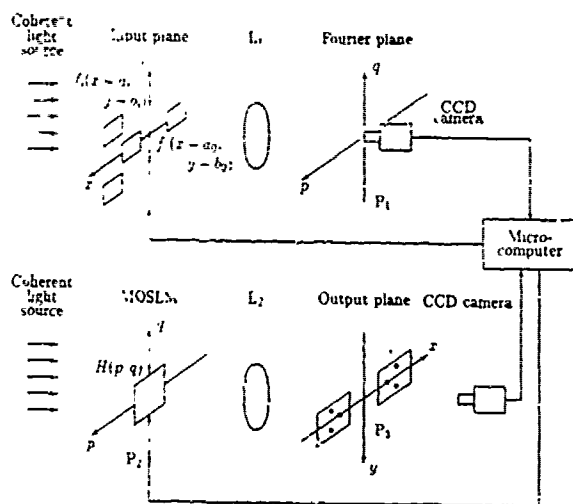$$F_i(p,q) = \mathcal{F}[f_i(x,y)], \quad i = 0,1,2,\dots,N, \qquad (2)$$

Fig. 1. Joint transform optical correlator using binarized interference fringes.

$\mathcal{F}$ denotes the Fourier transformation, and (p,q) is the angular spatial frequency coordinate system.

The corresponding intensity distribution captured by the CCD camera can be written as

$$|G(p,q)|^2 = \sum_{i=0}^{N} |F_i(p,q)|^2$$

$$+ \sum_{i=1}^{N} F_0(p,q)F_i^*(p,q)$$

$$\times \exp\{-j[(a_0 - a_i)p + (b_0 - b_i)q]\}$$

$$+ \sum_{i=1}^{N} F_0^*(p,q)F_i(p,q)$$

$$\times \exp\{-j[(-a_0 + a_i)p + (-b_0 + b_i)q]\}$$

$$+ \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} F_i(p,q)F_k^*(p,q)$$

$$\times \exp\{-j[(a_i - a_k)p + (b_i - b_k)q]\}$$

$$+ \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} F_i^*(p,q)F_k(p,q)$$

$$\times \exp\{-j[(-a_i + a_k)p + (-b_i + b_k)q]\}, \quad (3)$$

where $*$ represents the complex conjugate. The output light distribution, through inverse Fourier transformation, can be shown to be

$$g(x,y) = \sum_{i=0}^{N} R_{ii}(x,y)^2$$

$$+ \sum_{i=1}^{N} R_{0i}[x - (a_0 - a_i),y - (b_0 - b_i)q]$$

$$+ \sum_{i=1}^{N} R_{i0}[x + (a_0 - a_i),y + (b_0 - b_i)q]$$

$$+ \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} R_{ik}[x - (a_i - a_k),y - (b_i - b_k)q]$$

$$+ \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} R_{ki}[x + (a_i - a_k),y + (b_i - b_k)q], \quad (4)$$

where $R_{ih}(x,y)$ represents the cross correlation between $f_i(x,y)$ and $f_k(x,y)$. To separate the cross correlation between the reference object and the targets [i.e., $R_{0i}(x,y)$ and $R_{i0}(x,y)$] from the cross correlation between input targets themselves, we let the separation between the reference object and input targets be

$$\min_{i=1,\dots,N} [(a_0 - a_i)^2 + (b_0 - b_i)^2]$$

$$> \max_{i,k=1,\dots,N} [(a_i - a_k)^2 + (b_i - b_k)^2]. \quad (5)$$

If the input targets are identical to the reference object, i.e., $f_i(x,y) = f(x,y)$ for all i, Eq. (3) becomes

$$|G(p,q)|^2 = (N + 1)|F(p,q)|^2$$

$$+ 2 \sum_{i=1}^{N} |F(p,q)|^2 \cos[(a_0 - a_i)p + (b_0 - b_i)q]$$

$$+ 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} |F(p,q)|^2$$

$$\times \cos[(a_i - a_k)p + (b_i - b_k)q]. \quad (6)$$

Thus the carrier interference fringes can be seen in the cosine terms. By properly thresholding Eq. (6), the binarized fringe pattern can be described as

$$H(p,q) = \begin{cases} 1, & I(p,q) \geq I_T, \\ 0, & I(p,q) < I_T, \end{cases} \quad (7)$$
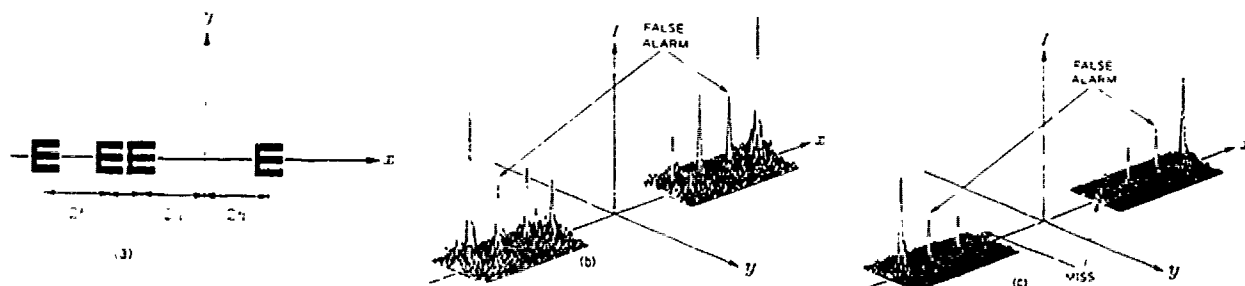

Fig. 2.   (a) Input plane geometry. E are used as the reference object and input targets. (b) Output correlation peaks after fringe binarization. (c) Output correlation peaks for noisy targets in (a) after fringe binarization.

where $I(p,q)$ is the intensity distribution, i.e., $|G(p,q)|^2$, and I_T is the selected threshold level.

Let us first consider a single input target case, then the binarized pattern $H(p,q)$ would become a rectangle wave function, which contains the carrier frequency (i.e., interference fringe frequency) and its harmonics. By inverse Fourier transforming this function, a series of correlation spots will be produced. The first-order correlation spots represent a match of the input targets with respect to the reference object. However, if the input is composed of multiple targets, the joint transform interference fringes will be very complicated and the binarization of these fringes could mask the carrier frequency content and possibly generate unwanted frequency components. Therefore, binarization of a joint power spectrum for multiple targets will generally cause false alarms and incorrect identification in the output correlation plane.

Computer simulations are now given to predict the effects of fringe binarization on the JTC. In the first example, shown in Fig. 2(a), the character E at the right-hand side is used as the reference object and the other characters on the left-hand side are used as input targets. From this configuration, the correlation between the reference and targets should be located inside a viewing window given by

$$3h < x < 8h \quad \text{and} \quad -h < y < h,$$

or

$$-8h < x < -3h \quad \text{and} \quad -h < y < h. \tag{8}$$

It is assumed that the character E occupies 15×15 pixels and the separation h is 17 pixels. A 512×512 point fast Fourier transform was used in the calculation. Figure 2(b) shows the output correlation peaks within the viewing window. A false alarm, due to the binarization process, can be seen in this simulation. In other words, if the nonlinear operation works well, there should be three correlation peaks in the viewing window. In fact, as shown in Fig. 2(b), there are four correlation peaks, one of which is apparently the result of the binarization process. We further demonstrate that, if the input targets [e.g., Fig. 2(a)] are embedded in additive Gaussian noise (with SNR = 15 dB), a false alarm and a miss are produced, as shown in the simulation result of Fig. 2(c). Moreover, if an A, instead of the E, is used as the reference object, as depicted in Fig. 3(a), a false alarm could also be produced, as shown in Fig. 3(b), even though the reference and targets are different. In this figure, we see a strong peak appear uninvitedly at a wrong position, which is purely due to the binarization process. Several runs of simulation, under low level noise environment (e.g., SNR > 15 dB), have been tested. The positions of the false alarm and the miss remain unaltered. Mention should be made that any change of the positions of the false alarm and the miss has not been observed in various runs of simulation, with different noise

disturbances of the same level. Thus we see that binarization of joint transform spectral fringes would generally introduce false alarms and misses for multitarget correlation.

Summarizing: we have examined the effects of the fringe binarization on the joint transform correlator via computer simulation. The simulated results show that the binarized joint transform interference fringes for multiple targets could produce false alarms and misses. Fringe binarization may not be usefully applied for multiple target recognition, although it can be applied successfully to noise free single target correlation.
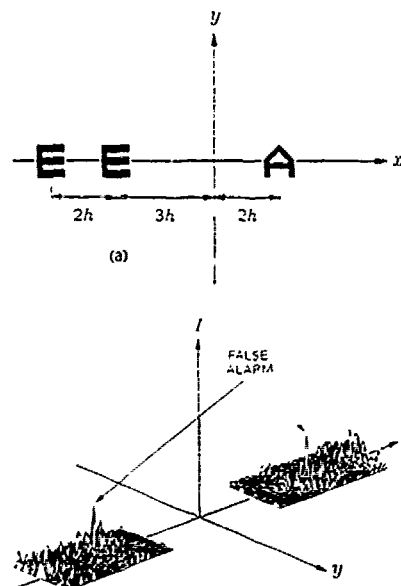
Fig. 3. (a) Input plane geometry. A is used as the reference object. (b) Output correlation peaks after fringe binarization.

References

1. C. S. Weaver and J. W. Goodman, "A Technique for Optically Convolving Two Functions," Appl. Opt. 5, 1248–1249 (1966).

2. J. E. Rau, "Detection of Differences in Real Distributions," J. Opt. Soc. Am. 56, 1490–1494 (1966).

3. F. T. S. Yu and J. E. Ludman, "Microcomputer-Based Programmable Optical Correlator for Automatic Pattern Recognition and Identification," Opt. Lett. 11, 395–397 (1986).

4. B. Javidi and S. F. Odeh, "Multiple Object Identification by Bipolar Joint Transform Correlation," Opt. Eng. 27, 295–300 (1988).

APPENDIX 10.3

Detection Efficiencies

Comparison of detection efficiencies for VanderLugt and joint transform correlators

Francis T. S. Yu, Q. W. Song, Y. S. Cheng, and Don A. Gregory

The correlation of peak intensities of VanderLugt and joint transform correlators with a single object, multiple objects, and noisy environments are analyzed. The study shows that the VanderLugt correlator can generally perform better for the multiple object case and also under a noisy environment.

I. Introduction

The VanderLugt correlator[1] (VLC) and the joint transform correlator[2] (JTC) are the major types of optical correlators used for pattern recognition applications.[3-9] In this paper, we investigate the detection efficiencies of these two types of correlator for multiple objects, and within a noisy environment.

II. Basic Analysis

In the following, we review the basic mathematics associated with the VanderLugt and Joint Transform correlators to establish notation. The complex spatial filter as obtained with the VanderLugt technique can be written as

$$H(f_x,f_y) = |F(f_x,f_y)|^2 + A^2 + 2A|F(f_x,f_y)| \times \cos[2\pi f_0 x - \phi(f_x,f_y)], \quad (1)$$

where $[f_x = x/(\lambda f), f_y = y/(\lambda f)]$ are the spatial frequency coordinates, f_0 is the spatial carrier frequency of the matched filter, A is the amplitude of the reference wave, λ is the illuminating wavelength, f is the focal length of the Fourier transform lens, and $F(f_x,f_y) = |F(f_x,f_y)| \exp[i\phi(f_x,f_y)]$ is the Fourier transform of the object $f(x,y)$.

The normalized amplitude transmittance of the spatial filter can be written as

$$H(f_x,f_y) = \frac{1}{[|F(0,0)| + A]^2} \{|F(f_x,f_y)|^2 + A^2 + 2A|F(f_x,f_y)| \cos[2\pi f_0 x - \phi(f_x,f_y)]\}, \quad (2)$$

where we have assumed

Don Greg· is with U.S. Army Missle Command. Redstone Arsenal, Alabama 35898; the other authors are with Pennsylvania State University, Electrical Engineering Department, University Park, Pennsylvania 16802.

$$|F(0,0)| = |F(f_x,f_y)|_{max}. \quad (3)$$

At the output plane, the complex light distribution can be written as

$$g(x,y) = \frac{1}{[|F(0,0)| + A]^2} \{f(x,y) * f(x,y) \otimes f(x,y) + A^2 f(x,y) \\ + A f(x + \lambda f f_0, y) \otimes f(x + \lambda f f_0, y) \\ + A f(x - \lambda f f_0, y) * f(x - \lambda f f_0, y)\}, \quad (4)$$

where $*$ and $\otimes$ represent the convolution and the correlation operation, respectively. The third term of the above equation represents the autocorrelation of $f(x,y)$, in which the correlation peak intensity is given by

$$I_V = \frac{A^2}{[|F(0,0)| + A]^4} \left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2 \, dx \, dy \right|^2. \quad (5)$$

For a given signal, the equation is a function of the reference beam amplitude. The optimum correlation peak intensity occurs for unity reference-to-object beam ratio (i.e., $A = |F(0,0)|$). The corresponding output peak intensity can be written as

$$I_V = \frac{1}{16|F(0,0)|^2} \left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2 \, dx \, dy \right|^2. \quad (6)$$

We shall now provide the same calculation for a joint transform correlator. Let the amplitude transmittance at the input plane be

$$t(x,y) = f(x - x_0, y) + f(x + x_0, y), \quad (7)$$

where one of the input objects serves as the reference image.

Since the square law detector at the Fourier plane takes the square modulus of the Fourier transform of $t(x,y)$, the normalized amplitude transmittance of the JTC filter can be written as

$$H(f_x,f_y) = \frac{|F(f_x,f_y)|^2}{2|F(0,0)|^2} [1 + \cos(4\pi x_0 f_x)]. \quad (8)$$

Assuming unity illumination of this filter, the output complex light distribution is given by

$$g(x,y) = \frac{1}{4|F(0,0)|^2} [2f(x,y) \otimes f(x,y)$$

$$+ f(x + 2x_0,y) \otimes f(x + 2x_0,y)$$

$$+ f(x - 2x_0,y) \otimes f(x - 2x_0,y)], \qquad (9)$$

in which the correlation peak intensity at $(2x_0,0)$ or $(-2x_0,0)$ is

$$I_J = \frac{1}{16|F(0,0)|^4} \left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2 \, dx \, dy \right|^2. \qquad (10)$$

From Eqs. (6) and (10), we have

$$\frac{I_V}{I_J} = |F(0,0)|^2. \qquad (11)$$

So far we have assumed unity read-out illumination for both the VLC and the JTC. If the read-out intensity for the JTC is B^2 instead of unity, then Eq. (11) would be equal to $|F(0,0)|^2/B^2$. Unless B is equal to $|F(0,0)|$, I_V would not equal I_J. This means for a smaller object with lower average transmittance, the ratio I_V/I_J is low. Thus, the relative performance would favor the JTC for smaller objects with lower average transmittance. An example can be given to demonstrate this effect. Let the average amplitude transmittance of the object $\bar{t}$ and its spatial dimension be X^2. Then, the zero-order Fourier diffraction would be confined within the region of $(\lambda f/X)^2$. Hence, the intensity at the origin of the Fourier transform plane is $\bar{t}^2 X^4/(\lambda f)^2$ for unity illumination. To make I_V smaller than I_J, for $\lambda = 6 \times 10^{-4}$ mm and $f = 300$ mm, $\bar{t}$ should be smaller than 0.18 for an object size of 1 mm².

III. Multiobject Detection

For the case of multiple objects, the amplitude transmittance of the VanderLugt matched filter, as given in Eq. (2), is independent of the input objects. The input amplitude transmittance is assumed to be:

$$t_{MV}(x,y) = \sum_{n=1}^{M} f(x - x_n,y), \quad x_1 < x_2 < x_3 \ldots < x_M, \qquad (12)$$

where we have M nonoverlapping objects distributed on the x-axis.

Thus, the output complex light distribution would be

$$g_{MV}(x,y) = \frac{1}{||F(0,0)| + A|^2} \sum_{n=1}^{M} [f(x - x_n,y)$$

$$\cdot f(x - x_n,y) \otimes f(x - x_n,y)$$

$$+ A^2 f(x - x_n,y) + Af(x + \lambda f f_0 - x_n,y)$$

$$\otimes f(x + \lambda f f_0 - x_n,y)$$

$$+ Af(x - \lambda f f_0 + x_n,y) \cdot f(x - \lambda f f_0 + x_n,y)], \qquad (13)$$

where the width of each object is assumed to be X. If $\lambda f f_0$ is greater than $x_M - x_1 + (5/2)X$ the outputs do not overlap. The corresponding correlation peak intensity at $(-\lambda f f_0 + x_n,0)$ is then identical to that of Eq. (6).

For the JTC, the input transmittance can be written as

$$t_{MJ} = f(x + x_0,y) + \sum_{n=1}^{M} f(x - x_n,y), \qquad (14)$$

where the first term represents the reference object.

After square law detection, the normalized amplitude transmittance of the JTC filter is

$$H_{MJ}(f_x,f_y) = \frac{|F(f_x,f_y)|^2}{(1 + M)^2 |F(0,0)|^2}$$

$$\times \left\{ 1 + M + 2 \sum_{n=1}^{M} \cos[2\pi(x_0 + x_n)f_x] \right.$$

$$\left. + 2 \sum_{\substack{m=1 \\ n<m}}^{M} \cos[2\pi(x_m - x_n)f_x] \right\}. \qquad (15)$$

The corresponding output light distribution can be shown to be

$$g_{MJ}(x,y) = \frac{1}{(1 + M)^2 |F(0,0)|^2} \left\{ (1 + M)f(x,y) \otimes f(x,y) \right.$$

$$+ \sum_{n=1}^{M} f(x - x_0 + x_n,y) \otimes f(x - x_0 + x_n,y)$$

$$+ \sum_{n=1}^{M} f(x + x_0 - x_n,y) \otimes f(x + x_0 + x_n,y)$$

$$\left. + \sum_{\substack{m=1 \\ n \neq m}}^{M} \sum_{n=1}^{M} f(x - x_m + x_n,y) \otimes f(x - x_m + x_n,y) \right\}, \qquad (16)$$

Thus, the correlation peak intensity at either $(x_0 - x_n,0)$ or $(-x_0 + x_n,0)$ can be written as

$$I_{MJ} = \frac{1}{(1 + M)^4 |F(0,0)|^4} \left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2 \, dx \, dy \right|^2. \qquad (17)$$

It is apparent that the intensity of the correlation peak decreases very rapidly as the number of the objects M increases. The reason is that the wave amplitude at the origin of the Fourier transform plane is proportional to the number of input objects, which includes the reference image. This is essentially limited by the maximum amplitude transmittance of the JTC filter, which can not exceed unity. Furthermore, from Eqs. (17) and (6), the correlation peak intensity ratio between the JTC and the VLC, for the multiobject case can be written as

$$\frac{I_{MJ}}{I_{MV}} = \frac{16}{(1 + M)^4 |F(0,0)|^2}. \qquad (18)$$

Since the output correlation intensity from the JTC is dependent on the number of input objects, Eq. (18) shows that the performance of the VLC would generally be better than the JTC, unless $(1 + M)^2 |F(0,0)| < 4$.

Figures 1(a)–(d) show the correlation peak intensity as a function of the number of input objects for different values of $|F(0,0)|^2$. Since the VanderLugt filter is independent of this input object illumination, the VLC peak intensity remains constant regardless of the num-
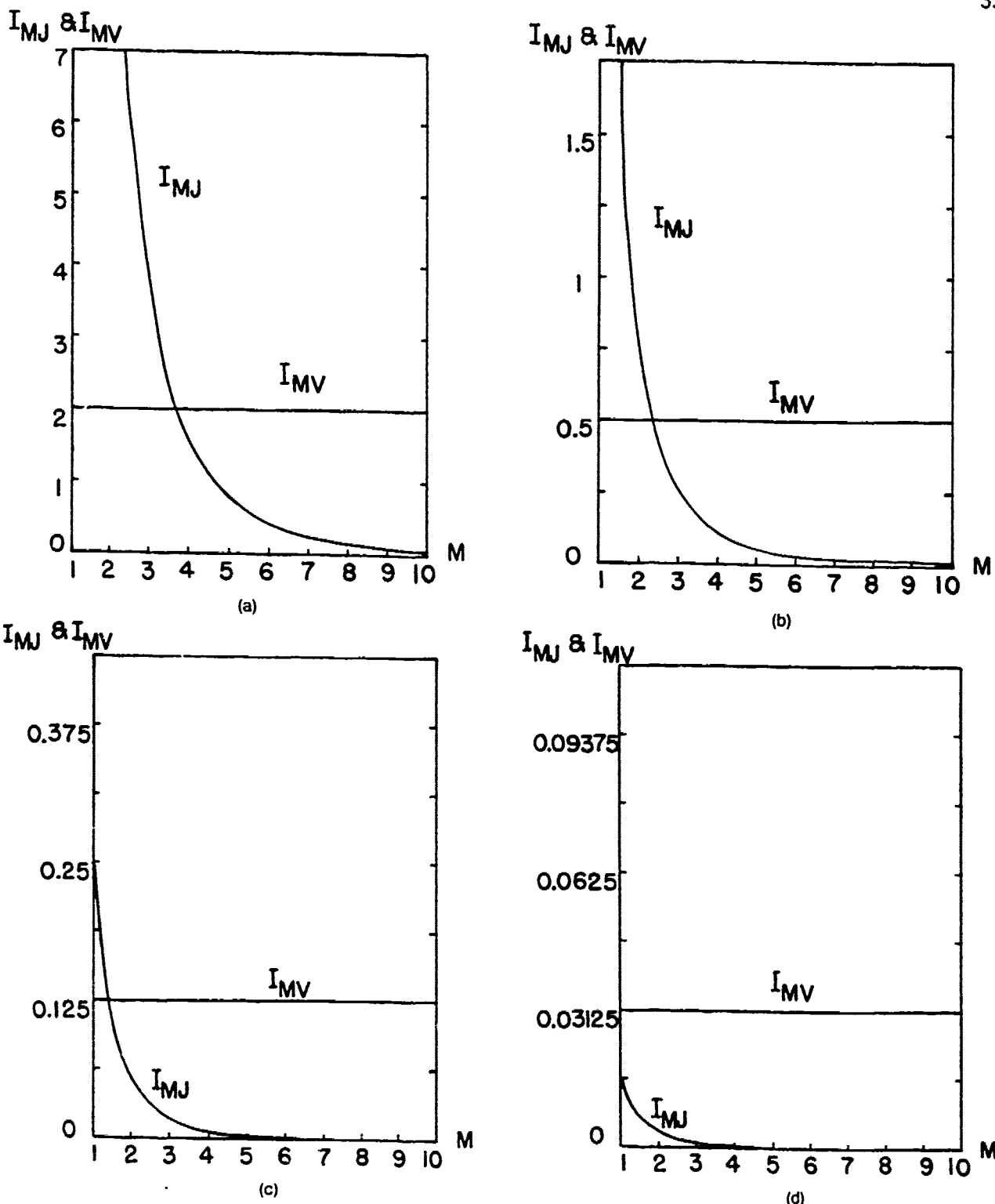
Fig. 1. Correlation peak intensity as a function of the number of objects M: I_{MJ} and I_{MV}, peak intensity of JTC and VLC, respectively; (a) for $|F(0,0)|^2 = 0.03125$; (b) for $|F(0,0)|^2 = 0.125$; (c) for $|F(0,0)|^2 = 0.5$; (d) for $|F(0,0)|^2 = 2.0$.

ber of input objects. However, the peak intensity decreases very rapidly for the JTC as the number of input objects increases. Furthermore, the relative performance of the JTC and VLC depends on the number of input objects for $|F(0,0)| < 1$. This can be demonstrated by looking at Fig. 3(a) in which the JTC performs better only when the number of input objects is smaller than three. However, if $|F(0,0)| > 1$, the VLC would always perform better than the JTC.

IV. Detection in a Noisy Environment

We shall now investigate the case for input objects imbedded in a noisy environment. In this case, the input transmittance function can be written as

$$t'_{MV}(x,y) = \sum_{n=1}^{M} f(x - x_n, y) + n(x,y), \qquad (19)$$

where $n(x,y)$ represents additive noise. Since the amplitude transmittance of a VanderLugt filter is independent of the input object, the complex light distribution at the output plane can be shown to be

$$\delta'_{MV}(x,y) = \frac{1}{[|F(0,0)| + A]^2} \left\{ \sum_{n=1}^{M} [f(x - x_n, y) * f(x - x_n, y)] \right.$$

$$\otimes f(x - x_n, y)$$

$$+ A^2 f(x - x_n, y) + A f(x + \lambda f f_0 - x_n, y)$$

$$\otimes f(x + \lambda f f_0 - x_n, y)$$

$$+ A f(x - \lambda f f_0 + x_n, y) * f(x - \lambda f f_0 + x_n, y)]$$

$$+ n(x,y) * f(x,y) \otimes f(x,y) + A^2 n(x,y)$$

$$+ A n(x + \lambda f f_0, y) \otimes f(x + \lambda f f_0 - x_n, y)$$

$$+ A n(x - \lambda f f_0, y) * f(x - \lambda f f_0 + x_n, y)] \right\}. \qquad (20)$$

Thus, the output correlation peak intensity is unchanged from that given in Eq. (6). The output noise intensity can be written as

$$I'_{MV} = \frac{1}{16|F(0,0)|^2} \left| \int\int_{-\infty}^{+\infty} n(x,y) f^*(x,y)\, dx\, dy \right|^2. \qquad (21)$$

The output signal-to-noise ratio can be expressed as

$$\left(\frac{S}{N}\right)_{VLC} = \frac{\left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2\, dx\, dy \right|^2}{\left| \int\int_{-\infty}^{+\infty} n(x,y) f^*(x,y)\, dx\, dy \right|^2}. \qquad (22)$$

For a noisy input in a JTC, the input is given by

$$t'_{MJ}(x,y) = f(x - x_0, y) + \sum_{n=1}^{M} f(x - x_n, y) + n(x,y). \qquad (23)$$

The amplitude transmittance of the JTC filter can be written as

$$H_{MJ}(f_x, f_y) = \frac{1}{[(1 + M)|F(0,0)| + |N(0,0)|]^2} \left[|F(f_x, f_y)|^2 \left\{ 1 + M \right.\right.$$

$$+ 2 \sum_{n=1}^{M} \cos[2\pi(x_0 - x_n) f_x]$$

$$+ 2 \sum_{n=1}^{M} \sum_{\substack{n=1 \\ n<m}}^{M} \cos[2\pi(x_m - x_n) f_x] \right\}$$

$$+ F(f_x, f_y) N^*(f_x, f_y) \exp[-i2\pi x_0 f_x]$$

$$+ F^*(f_x, f_y) N(f_x, f_y) \exp[i2\pi x_0 f_x]$$

$$+ F(f_x, f_y) N^*(f_x, f_y) \sum_{n=1}^{M} \exp[-i2\pi x_n f_x]$$

$$+ F^*(f_x, f_y) N(f_x, f_y)$$

$$\times \sum_{n=1}^{M} \exp[i2\pi x_n f_x] + |N(f_x, f_y)|^2 \right]. \qquad (24)$$

The output correlation peak and noise intensity can be shown to be, respectively,

$$I'_{MJ} = \frac{1}{[(1 + M)|F(0,0)| + |N(0,0)|]^4}$$

$$\times \left| \int\int_{-\infty}^{+\infty} |f(x,y)|^2\, dx\, dy \right|^2, \qquad (25)$$

$$I'_{MJ} = \frac{1}{[(1 + M)|F(0,0)| + |N(0,0)|]^4}$$

$$\times \left| \int\int_{-\infty}^{+\infty} n(x,y) f^*(x,y)\, dx\, dy \right|^2. \qquad (26)$$

From these results, it is observed that the output S/N is the same for both the JTC and the VLC. However, the correlation peak intensities are quite different. For the JTC, the correlation intensity is dependent on the presence of input noise, while for the VLC, the peak intensity is independent. The ratio of the peak intensity between these correlations can be written as:

$$\frac{I'_{MJ}}{I'_{MV}} = \frac{16|F(0,0)|^2}{[(1 + M)|F(0,0)| + |N(0,0)|]^4}. \qquad (27)$$

Figure 2(a) shows logarithmic plots of I'_{MJ}/I'_{MV} as a function of M, for various noise levels $|N(0,0)|^2/|F(0,0)|^2$. For a certain noise level, $ln(I'_{MJ}/I'_{MV})$ decreases monotonically as M increases. On the other hand, for a fixed number of input objects, the value of $ln(I'_{MJ}/I'_{MV})$ becomes smaller as the noise level increases. In the region above the M-axis, the correlation peak intensity for the JTC is higher than that for the VLC. In other words, the JTC would perform better in this region. On the other hand, the VLC would perform better than the JTC in the region below the M-axis. Furthermore, the curves in Fig. 2(a) are shifted down as $|F(0,0)|^2$ increased. This implies that due to the limited dynamic range of the square law converter, the JTC correlation intensity decreases if the average transmittance function of the object becomes larger. Logarithmic correlation-peak-intensity ratios as a function of noise level for various number of input objects are plotted in Fig. 3. It is apparent that $ln(I'_{MJ}/I'_{MV})$ decreases monotonically as $|N(0,0)|^2/|F(0,0)|^2$ increases. With the same noise level, the value of the correlation intensity ratio becomes lower as the number of input objects increases.

V. Conclusion

We have studied the detection efficiencies for the JTC and the VLC, with multiobject scenes and in a noisy environment. For the single-object case. it was found that the relative performance. between these two correlators, depends upon the object transmittance function and the object size. In general, the VLC would perform better unless the input object is very small with very low average transmittance.

For the multi-object case, the JTC peak intensity decreases rapidly as the number of input objects increases. Since the VanderLugt filter is independent of
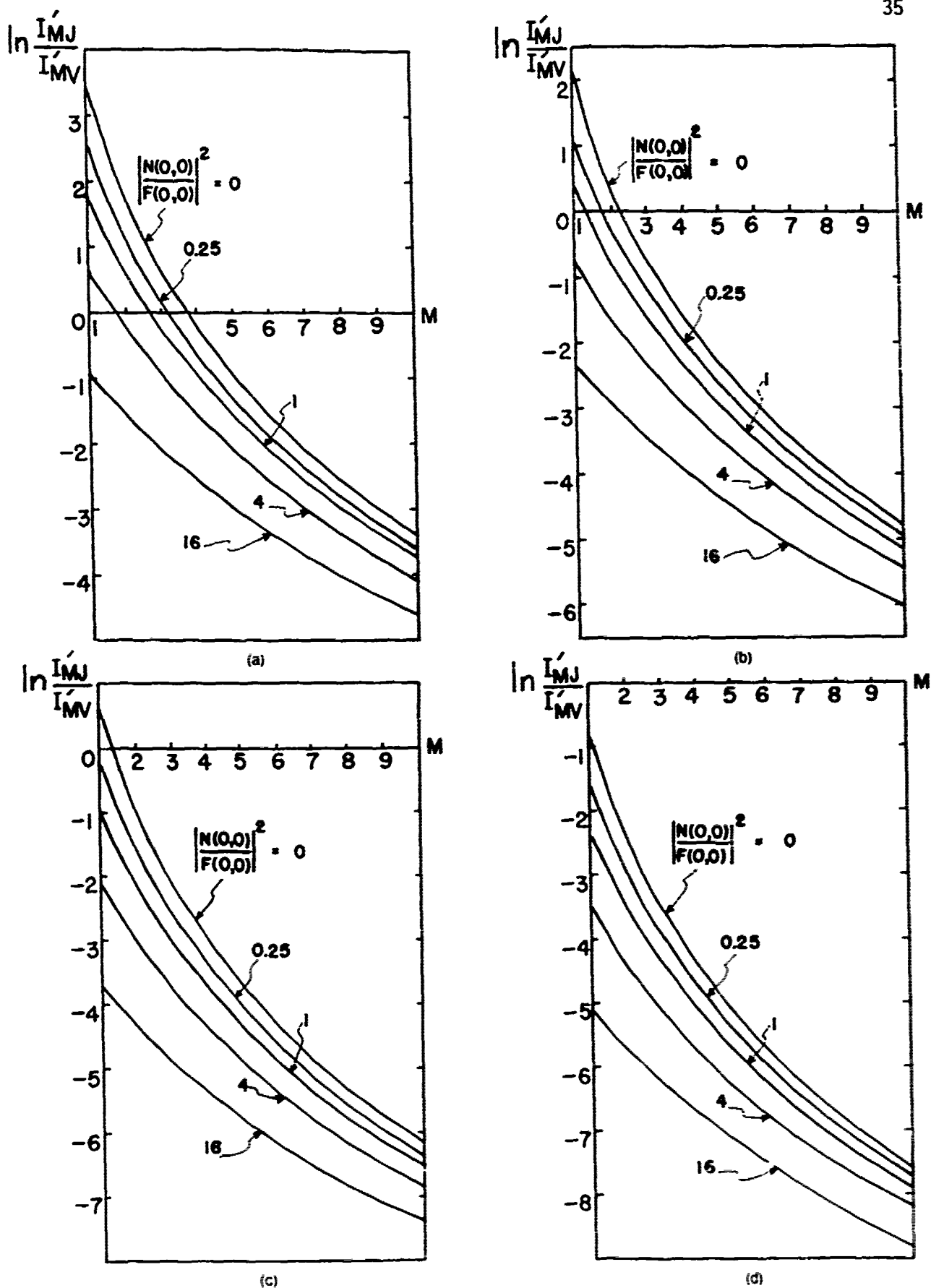
Fig. 2. Logarithmic intensity ratio $ln(I_{MJ}'/I_{MV}')$ as a function of number of objects M. (a) for $|F(0,0)|^2 = 0.03125$; (b) for $|F(0,0)|^2 = 0.125$; (c) for $|F(0,0)|^2 = 0.5$; (d) for $|F(0,0)|^2 = 2.0$.
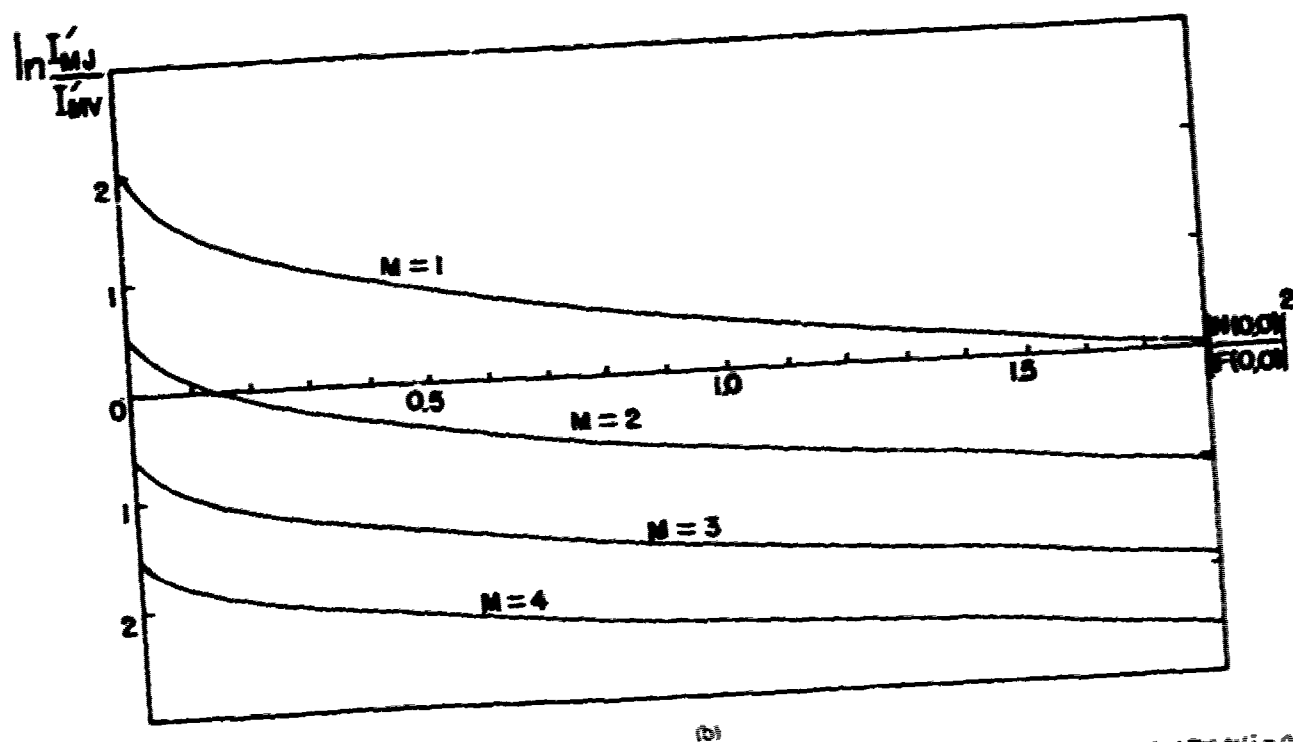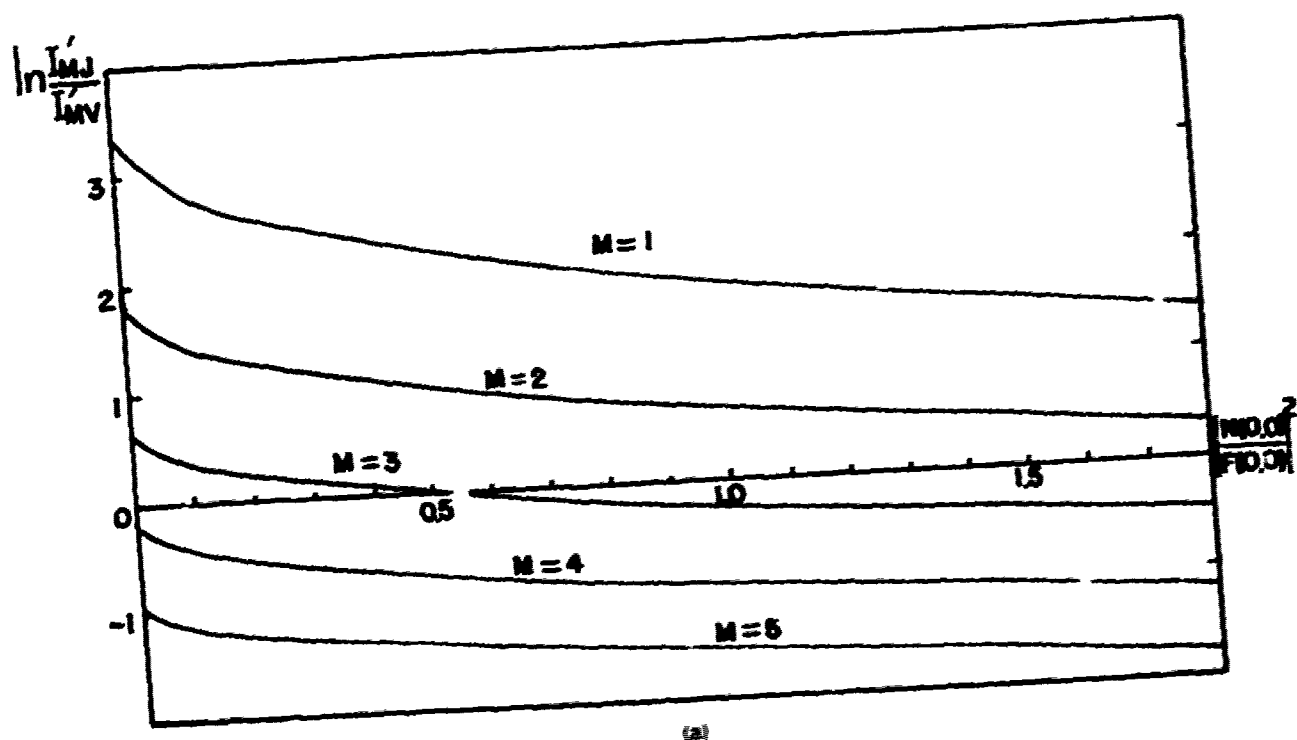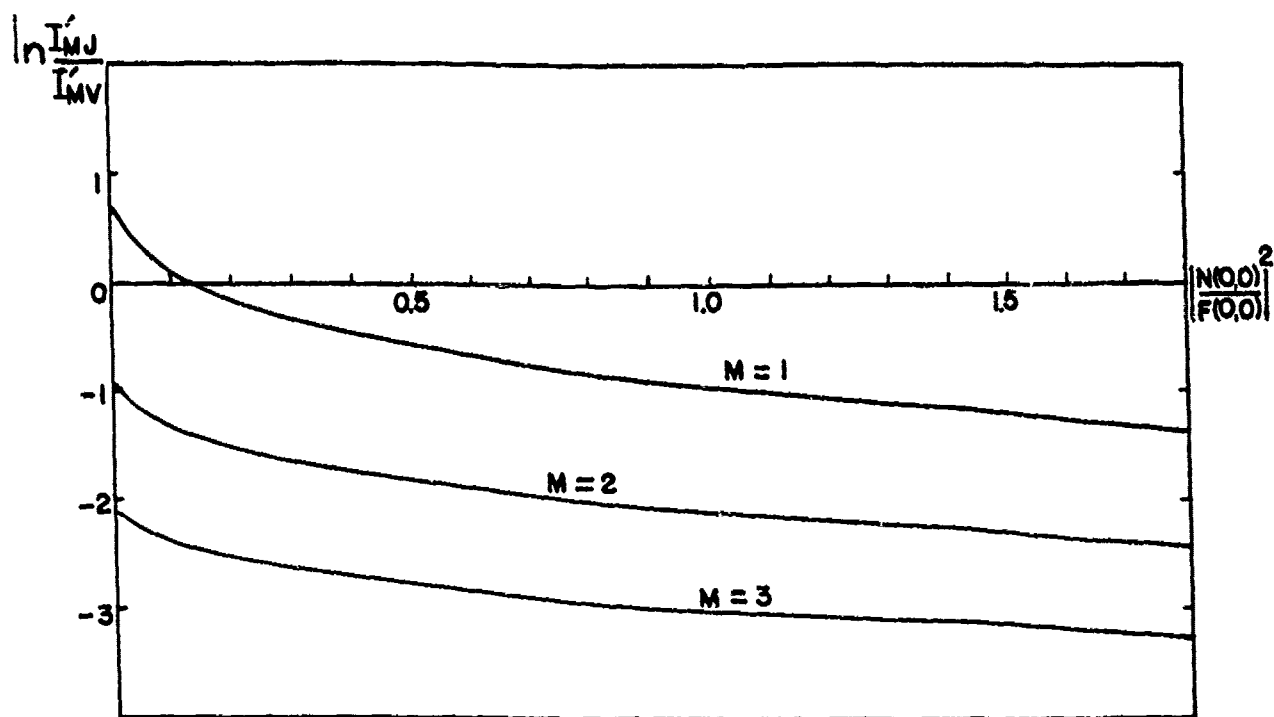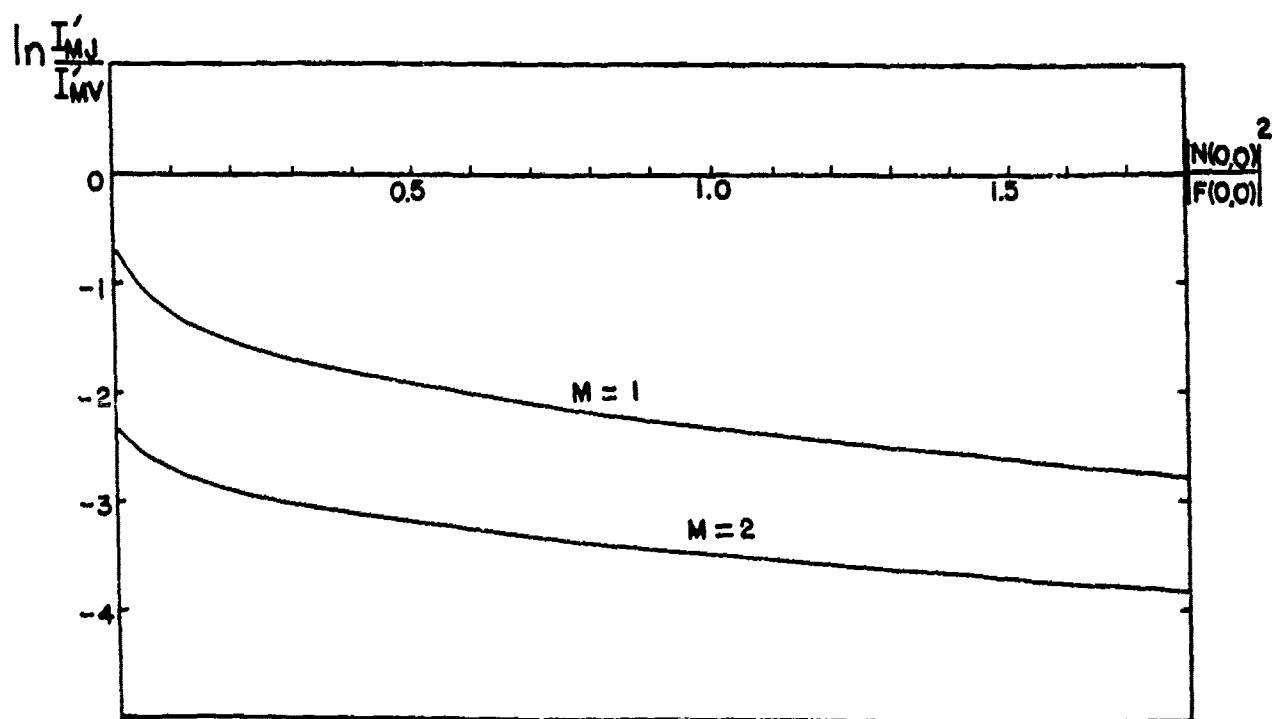
(a)

(b)

Fig. 3. (a) (b)   Logarithmic intensity ratio $\ln(I_{MJ}/I_{VJ})$ as a function of $|N(0,0)|^2/|F(0,0)|^2$.   (a) for $|F(0,0)|^2 = 0.03125$; (b) for $|F(0,0)|^2 = 0.125$.

(c)

(d)

Fig. 3. (c) (d)   Logarithmic intensity ratio $ln(I_{MJ}/(I_{VJ})$ as a function of $|N(0,0)^2/|F(0,0)|^2$:   (c) for $|F(0,0)|^2 = 0.5$; (d) for $|F(0,0)|^2 = 2.0$.

input noise, the VLC peak intensity remains constant regardless the number of input objects. When $|F(0,0)| < 1$, the relative performance depends on the number of input objects. The JTC may perform better if $(1 + M)^2|F(0,0)| < 4$. However, for $|F(0,0)|^2 > 1$, the VLC would always perform better than the JTC.

Although the output signal-to-noise ratios for both correlators are the same, their correlation peak intensities are quite different. Since the VanderLugt filter is not changed by input noise, it should generally perform better in a noisy environment, except for $|F(0,0)|[1 + M + |N(0,0)|/|F(0,0)|]^2 < 4$.

We finally remark that, in terms of detection efficiency, the VLC will generally perform better when compared to the JTC. However, with respect to real-time applications, the JTC still offers the advantages in spatial filter synthesis, alignment of spatial filter, and overall simplicity.

Y. S. Cheng is on leave from Institute of Optical Sciences, National Central University, Taiwan.

References

1. A. VanderLugt, "Signal Detection by Complex Spatial Filtering," IEEE Trans. Inf. Theory IT-10, 139–145 (1964).
2. C. S. Weaver and J. W. Goodman, "Technique for Optically Convolving Two Functions," Appl. Opt. 5, 1248–1249 (1966).
3. F. T. S. Yu and X. J. Lu, "A Real-time Programmable Joint Transform Correlator-to," Opt. Commun. 52, 10–16 (1984).
4. D. Casasent, "Coherent Optical Pattern Recognition: A Review," Opt. Eng. 24, 26–32 (1985).
5. F. T. S. Yu and X. J. Lu, "Large-Capacity Real-Time Spatial-Spectral Scanning Optical Correlator," Appl. Opt. 25, 812–816 (1986).
6. D. J. Ennis and D. A. Jared, "Optical Processing and Space Station Automation," Opt. Eng. 25, 808–820 (1986).
7. J. L. Horner and P. D. Gianino, "Phase-Only Matched filtering," Appl. Opt. 23, 812–816 (1984).
8. D. Psaltis, E. G. Peak, and S. S. Venkatesh, "Optical Image Correlation with a Binary Spatial Light Modulator," Opt. Eng. 23, 698–704 (1984).
9. D. A. Gregory, "Real-Time Pattern Recognition Using a Modified Liquid Crystal Television in a Coherent Optical Correlator," Appl. Opt. 25, 467–469 (1986).

APPENDIX 10.4

JTC Color Pattern Recognition

Adaptive joint transform correlator for real-time colour pattern recognition

F.T.S. YU, S. JUTAMULIA*, R.V. YELAMARTY, D. A. GREGORY

Colour liquid crystal televisions (LCTVs) can be used to perform real-time multi-channel joint transform correlation capable of discriminating between different colours and different shapes. In the future this optical system may be combined with a digital knowledge-based inference processor. Preliminary optical experimental results are presented, which demonstrate how such a system would work.

KEYWORDS: liquid crystal televisions, colour recognition, shape recognition, joint transform correlation

Introduction

Recently, we have demonstrated real-time polychromatic pattern recognition using a commercially available inexpensive colour liquid crystal television (LCTV) and multiple Vander Lugt matched spatial filters[1]. This method has the important limitation of changing the reference pattern. To improve the adaptability of the colour pattern recognition system, a joint transform correlation processor which has been successfully applied to monochromatic pattern recognition[2], is proposed. The reference colour pattern then may simply be generated by a computer or a CCD camera.

The uses of LCTVs in optical processing and computing have been widely reported[1-14]. Monochromatic real-time pattern recognition based on Vander Lugt matched spatial filtering techniques was demonstrated by Liu et al[4] and Gregory[5]. A non-coherent pattern recognition technique using LCTVs was reported by Jutamulia et al[14], which can be directly extended to colour object identification. Colour pattern recognition has also been performed using a colour sensitive matched spatial filter[15-17], a deflecting prism[18], and a diffraction grating[19-21]. In this paper, a technique which uses colour LCTVs in the construction of a multi-channel joint transform correlator will be described. This system could be considered as a model for future, optical, knowledge-based pattern recognition systems[22].

Concept

The joint transform processor is an alternative

FTSY and RVY are in the Electrical Engineering Department, Pennsylvania State University, University Park, Pennsylvania 16802, USA DAG is at the US Army Missile Command, Research, Development and Engineering Center, Research Directorate, Redstone Arsenal, Alabama 36898-5248, USA When the work was done SJ was at the Electrical Engineering Department, Pennsylvania State University; SJ is presently at the Quantex Corporation 2 Research Court, Rockville, Maryland 20850, USA Received 19 September 1988 Revised 8 February 1989

approach to optical pattern recognition[23-24], and has proven to be suitable for adaptive programmable correlation because no matched filter is required. In general, it is very difficult to generate a dynamic Vander Lugt matched spatial filter in the Fourier plane.

A commercially available colour LCTV (Radio Shack, Catalogue Number 16-154) has been used to construct the multi-channel joint transform correlator. The structure of the display panel is depicted in Fig. 1a. From this a diffraction pattern with well separated colour spectra should be expected as shown in Fig. 1b. Fig. 2 shows a schematic diagram of the experimental arrangement. The lower optical set-up in Fig. 2 performs the joint Fourier transform of the reference and test object displayed on LCTV1 and LCTV2 respectively. The upper one performs another Fourier transform operation on the joint transform power spectrum to obtain the correlation signal. L_1 and L_2 are lenses for collimating the beams. L_3 is a Fourier transform lens and L_4 is a microscope objective.

Basically this implementation utilizes three colour LCTVs. illuminated by collimated red and green laser light from HeNe and Ar ion lasers as shown in Fig. 2. Ideally, three primary colours should be used. Vidicon3 is a TV camera directed toward a scene which may or may not contain the desired colour object. This is displayed on LCTV1 without passing through the microcomputer or other digital system preserving the high speed processing capability. LCTV2 is driven by a computer in which a set of reference colour patterns have been stored. The reference pattern displayed on LCTV2 can be chosen directly by an operator or determined by software control of the feedback input from Vidicon2.

LCTV1 and LCTV2 are arranged side by side in the same plane as shown in the lower set-up of Fig. 2. The

Fig. 2 Conceptual architecture of an adaptive joint transform correlator

Fig. 3 Schematic diagram of a knowledge-based real-time polychromatic correlator
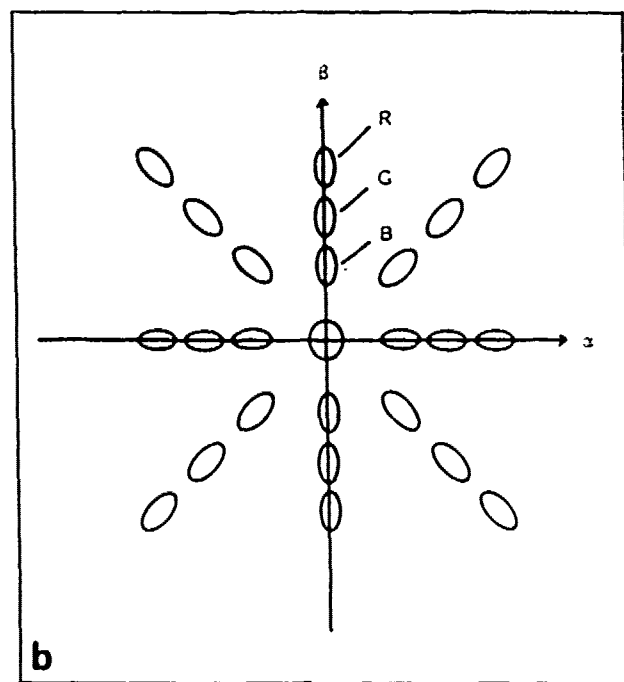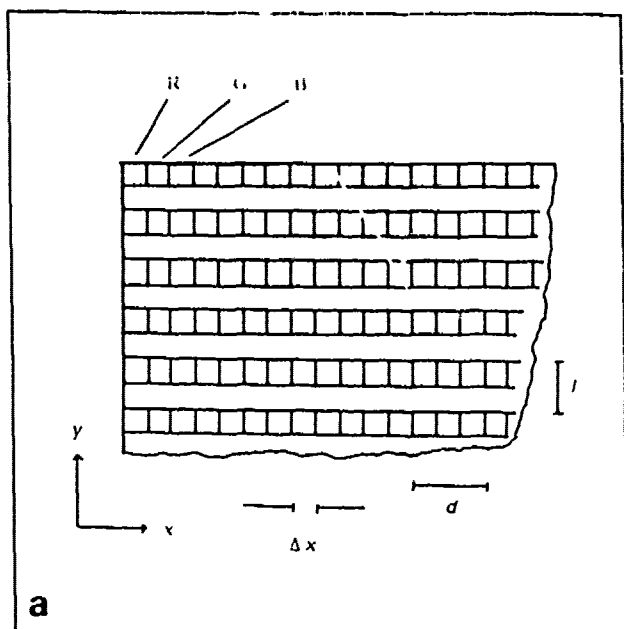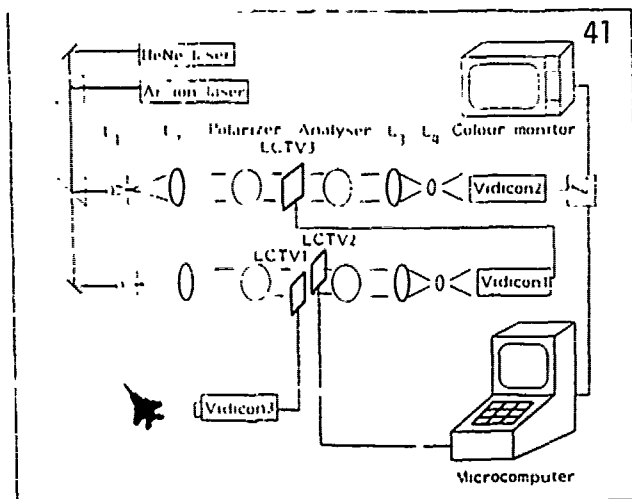
Fig. 1 (a) Structure of the colour LCTV display panel; (b) schematic diagram of the polychromatic spectral distribution at the Fourier transform plane (zero and first order maxima)

superimpose. Thus. a correlation peak which has the same colour as the object (if three primary colours are used) can be observed. In approaching a knowledge-based correlation system[22], a diffraction order that gives a spatially separated primary colour correlation pattern should be selected. A computer would analyse the independent colour signals and then determine the colour of the object. This may require updating the reference pattern with patterns from the knowledge base for obtaining the best match. A schematic diagram of this process is depicted in Fig. 3.

Experiments

To verify operation of this knowledge-based colour correlator and the usefulness of the colour LCTVs. a preliminary experiment has been performed. A schematic diagram of the experimental set-up is shown in Fig. 4. A microcomputer was used to generate a colour object and reference patterns simultaneously on a single LCTV. A black and white picture of the generated colour patterns (simulated fighters) taken from a colour video monitor is given in Fig. 5a. Fig. 5b shows the picture as taken directly from the LCTV and as can be seen. the LCTV has low resolution and give a colour ghost image. The colour ghost image is due to cross-talk between neighbouring colour pixels. For simplicity, only two primary colours (red and green) were used. A collimated bichromatic coherent beam from HeNe and Ar ion lasers is incident on the LCTV, which was disassembled and immersed in a
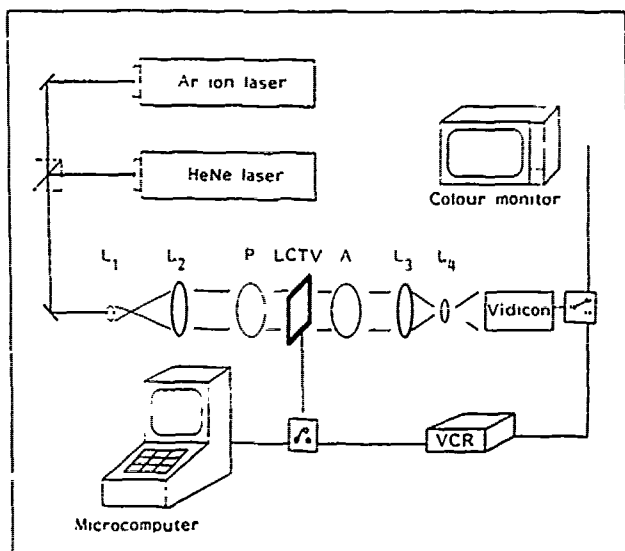
joint transform is performed by the lower lens L_3. The pixel structure of the LCTV produces multiple colour spectra in the Fourier plane as shown schematically in Fig. 1b. A microscope objective lens L_4 with a small aperture selects and magnifies only one diffraction order consisting of well separated colour spectra. The intensity joint transform distribution of the selected diffraction order can then be recorded by Vidicon1 which is directly connected to LCTV3.

The correlation signal is obtained by performing a Fourier transform of the intensity pattern of the colour joint transform spectra displayed on LCTV3. This is done using lens L_3. lens L_4 is another microscope objective which selects and magnifies the colour correlation signal for detection by Vidicon2. If the zero order is selected. colour correlation peaks will

Fig. 4 Experimental set-up of the adaptive joint transform correla₁or for real-time colour pattern recognition

liquid gate. The polarizers of the LCTV were removed and good quality polarizers were used in the set-up shown in Fig. 4. The joint transform of the object and the reference patterns is performed by lens L_3. If the object and reference are identical (same shape and colour). the joint transform spectrum consists of a two-colour fringe structure.

The pixel structure of the display panel diffracted the joint transform spectrum into many orders. The red and green spectra were overlapping at the zero order. However. they were separated further at higher diffraction orders. The second order of the joint transform spectrum was selected because the red and green spectra were sufficiently separated but they were still covered and could be magnified by a single microscope objective L_4. These magnified independent red and green spectra were detected by a TV camera and recorded by a video cassette recorder (VCR). The joint transform spectra displayed on a video monitor are shown in Fig. 6a for the input of Fig. 5b. Fig. 6a shows a black and white picture of green fringes on the left and red fringes on the right. The green fringes seem to be stronger than the red. This might be caused by the larger green area than the red as shown in Figs. 5a and b. When one of the simulated fighters was covered. the fringes disappeared as illustrated in Fig. 6b. The red spectrum was under-exposed thus it is hardly seen in Fig. 6b.

The recorded fringe structure was then replayed using the same LCTV. To reduce colour cross-talk (transmittance of red light through green pixels and vice versa). additional red and green filters were used to cover red and green fringes of the LCTV. The correlation signals were then input to the computer for the inference process. In this preliminary demonstration. a digital inference process was not attempted. The correlation outputs were obtained by taking the Fourier transform of the fringe structure shown in Fig. 6a Half of the magnified bichromatic correlation output is depicted in Fig. 7a. The extended dc patterns are shown in the upper part of Fig. 7a. The dc patterns were due to the background transmittance of the low contrast LCTV screen. The
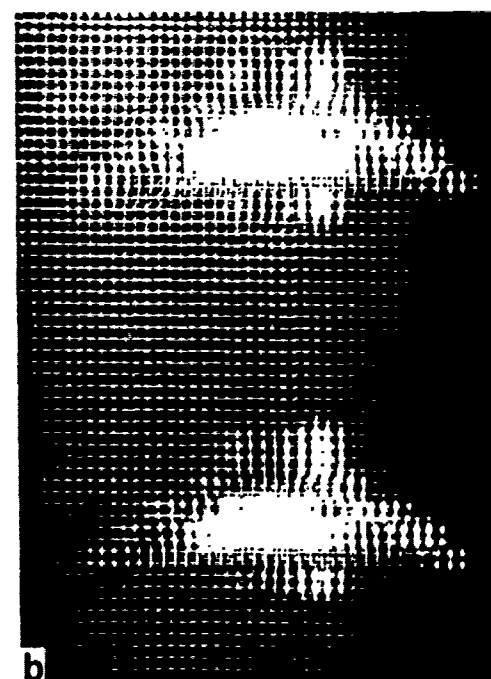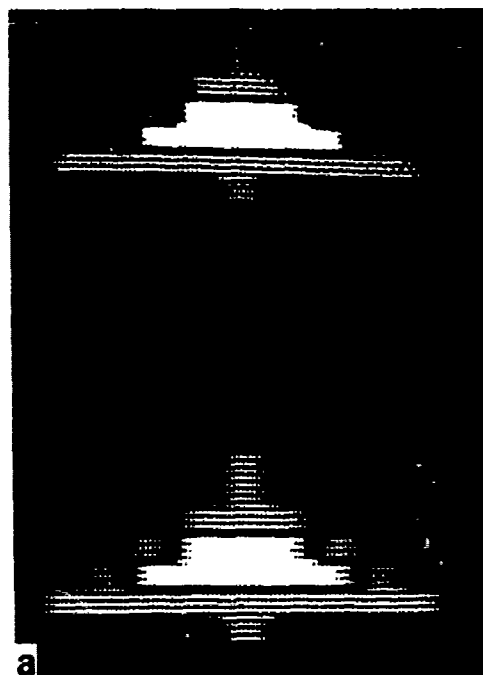

Fig 5 Object and reference patterns generated by microcomputer (Apple II) (a) displayed on ordinary colour TV montior. and (b) displayed on colour LCTV. The centre is red and the edge is green

Fourier transform of Fig. 6b is depicted in Fig. 7b. The subtraction of the dc struture shown in Fig. 7b from Fig. 7a will produce the correlation spot alone. The subtraction may be included in the digital inference process.

Concluding remarks

The application of a colour LCTV to polychromatic pattern recognition has been demonstrated using a multi-channel joint transform correlator. This should be suitable for extension to a knowledge-based object identifier because generation of the colour reference pattern in the joint transform correlator is easier than
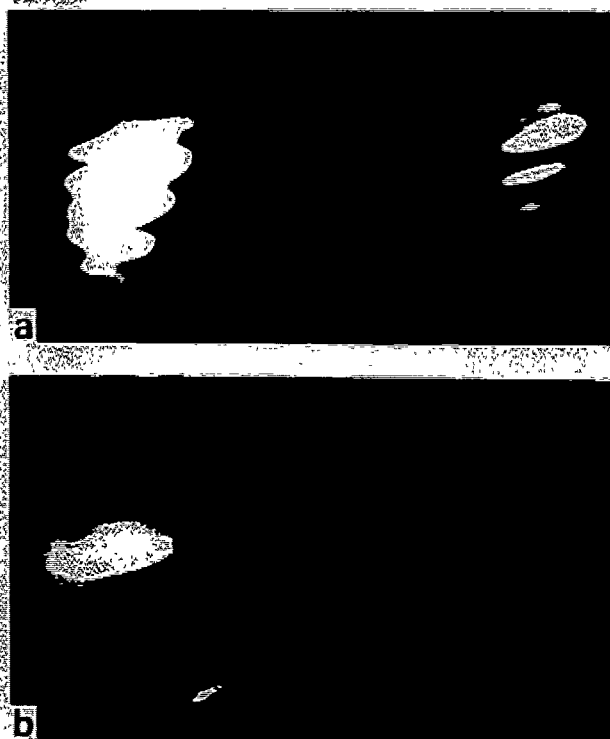
43

Fig. 6 (a) Joint transform power spectra of Fig. 5. the left and the right fringe structures are green and red respectively; (b) fringe structures disappear when one fighter in Fig. 5 is covered (the red spectrum was underexposed)
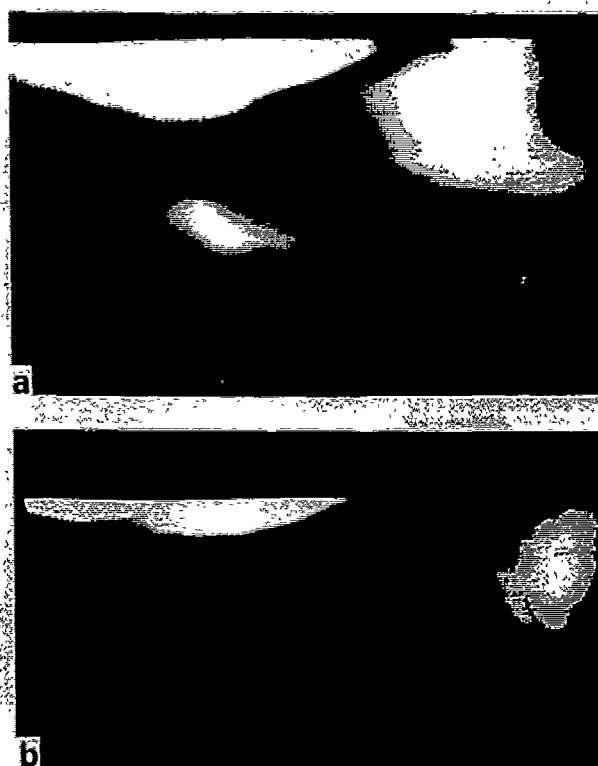

Fig. 7 (a) Fourier transform of Fig. 6 (a). The small spots at the bottom indicate the correlation terms; (b) fourier transform of a dark screen displayed on the LCTV, showing the dc terms. The subtraction of (b) from (a) will give the correlation terms alone. The left and the right structures are green and red respectively

that of the traditional Vander Lugt correlator. The system is capable of discriminating between different colours and different shapes. Although LCTV possesses limited space-bandwidth product (SBP). it is a low-cost. computer-addressable. spatial light modulator capable of processing polychromatic signals. With further improvements. the colour LCTV may be applicable to colour-sensitive knowledge-based pattern recognition.

Acknowledgement

References

1 Yu. F.T.S., Jutamulia. S., Lin. T.W. 'Real-time polychromatic signal detection using a colour liquid crystal television'. Opt Engng. 26. (1987) 453

2 Yu. F.T.S., Jutamulia. S., Lin. T.W.. Gregory. D.A. 'Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator'. Appl Opt. 26. (1987) 1370

3 McEwan, J.A.. Fisher. A.D.. Rolsma. P.B.. Lee. J.N. 'Optical processing characteristic of a low-cost liquid crystal display device'. J Opt Soc Am. A2(13). (1985) 8

4 Liu. H.K.. Davis. J.A.. Lilly. R.A. 'Optical data processing properties of a liquid crystal television spatial light modulator'. Opt Lett. 10. (1985) 635

5 Gregory. D.A. 'Real-time pattern recognition using a modified liquid crystal television in a coherent optical correlator'. Appl Opt. 25. (1986) 469

6 Young, M. 'Low-cost LCD video display for optical processing'. Appl Opt. 25. (1986) 1024

7 Tai. A.M. 'Low-cost LCD spatial light modulator with high optical quality'. Appl Opt. 25. (1986) 1380

8 Casasent. D.P.. Xia. S.X. 'Phase correction of light modulators'. Opt Lett. 11 (1986) 398

9 Yu. F.T.S.. Jutamulia. S.. Huang. H.L. 'Experimental applications of low-cost liquid crystal television to white-light optical signal processing'. Appl Opt. 25. (1986) 3324

10 Mok. F.. Diep. J.. Liu. H.K.. Psaltis. D. 'Real-time computer-generated hologram by means of liquid crystal television spatial light modulator'. Opt Lett. 11. (1986) 748

11 Hughes. K.D.. Rogers. S.K.. Mills. J.P.. Kabrisky. M. 'Optical processing using liquid crystal televisions'. Appl Opt. 26. (1987) 1042

12 Yu. F.T.S.. Jutamulia. S.. Gregory. D.A. 'Real-time liquid crystal TV XOR-and-XNOR gate binary image subtraction technique'. Appl Opt. 26. (1987) 2738

13 Yu. F.T.S.. Jutamulia. S.. Gregory. D.A. 'Optical parallel logic gates using inexpensive liquid crystal televisions'. Opt Lett. 12. (1987) 1050

14 Jutamulia. S.. Lin. T.W.. Yu. F.T.S. 'Real-time noncoherent correlator using liquid crystal television'. Opt Commun. 64. (1987) 115

15 Shi. N.K. 'Color-sensitive spatial filters'. Opt Lett. 3. (1978) 85

16 Case. S.K. 'Pattern recognition with wavelength multiplexed filters'. Appl Opt. 18. (1979) 1890

17 Ishii. Y.. Murata. K. 'Color-coded character recognition experiment with wavelength triplexed. reflection-type holographic filters'. Opt Lett. 7. (1982) 230

18 Braunnecker. B.. Bryngdahl. O. 'Multiplex optical processing combination of polychromatic light and a dispersive element'. Opt Commun. 40. (1982) 332

19 Yu. F.T.S.. Chao. T.H. 'Color signal correlation detection by matched spatial filtering'. Appl Phys. B:32. (1983) 1

20 Yu. F.T.S.. Javidi. B. 'Experiments on real-time polychromatic signal detection by matched spatial filtering'. Opt Commun. 56. (1986) 334

21 Yu. F.T.S.. Lu. X. 'Large capacity real-time spatial spectral scanning optical correlator'. Appl Opt. 25. (1986) 812

22 Casasent. D.P.. Liebowitz. B. 'Model based knowledge-based optical processor'. Appl Opt. 26. (1987) 1935

23 Lee. S.H. 'Coherent optical processing'. in Optical Information Processing. Ed. S.H. Lee. Springer-Verlag. New York. (1981) 61

24 Yu. F.T.S.. Lu. X.J. 'A real-time programmable joint transform correlator'. Opt Commun. 52. (1984) 10

APPENDIX 10.5

Comparison Vander Lugt and Joint Transform Correlators

Comparison of Vander Lugt and Joint Transform Correlators

X. J. Lu, F. T. S. Yu, and Don A Gregory

Department of Electrical Engineering, University Park, P 16802. USA
U.S. Army Missile Command, Redstone Arsenal, AL 35898-5248, USA

Abstract. A generalized optical correlator is used to analyze the performances of the Vander Lugt and the joint transform architectures. Noise performance and the effects caused by the differences between the input and the reference functions are analyzed. It is found that the modulation index of the joint transform filter decreases due to the presence of noise as well as the differences between the input and the reference functions. However, it is found that, in general, the joint transform correlator would perform better than the Vander Lugt correlator and certainly has the advantage for real-time implementation.

PACS: 42.30. 42.80

Two commonly used techniques for optical pattern recognition are the Vander Lugt [1] and the joint transform [2, 3] correlators. With the Vander Lugt correlator, investigators have used a variety of techniques to synthesize multiplexing matched filters, such as the frequency [4], the angular [5], the color multiplexing [6], the holographic lens array [7] and the rotating grating technique [8]. However, the joint transform architecture combined with a TV camera such as vidicon and a CCD can be operated in a real-time mode with ease [9, 10]. Nevertheless, the application of spatial light modulators to both Vander Lugt and joint transform correlators provide significant improvement of the system operation [8–13]. We shall, in this paper, analyze the basic performance of the Vander Lugt (VLC) and the joint transform (JTC) correlators. In the following, we first present a generalized optical correlator for both architectures. The performance and limitations of these correlators are evaluated. Multiplicative noise introduced by the spatial light modulators has been used for the analyses. The effects of the noise perturbation and the differences between the input and the reference functions on modulation index of the correlation filter are calcu ~d. It is shown that the JTC can generally perform better than the VLC. particularly for real-time implementation.

1. A Generalized Optical Correlator

A generalized optical correlator can be described by dividing the system performance into two parts: namely, the correlation filter construction and the correlation operation. Notice that the term, "correlation filter," used here is for both the matched filter (MF) and the joint transform hologram (JTH). The construction of these filters is basically using the same procedure as shown in Fig. 1. where an input transparency $h(x, y-b)$ and a reference function $g(x, y+b)$ at the input plane P_1 are illuminated by a collimated coherent light beam with wavelength λ. The intensity distribution of the joint Fourier spectrum at P_2 is given by

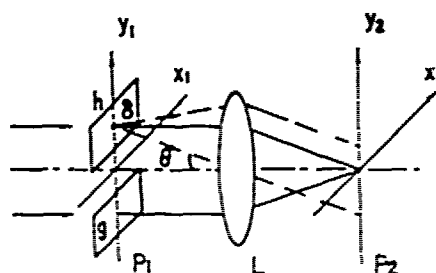$$I(u, v) = |H|^2 + |G|^2 + HG^* e^{-i4\pi v b} + H^* G e^{i4\pi v b}. \qquad (1)$$

Fig. 1. Construction of a correlation filter

where $2b$ is the mean separation of the two functions and u and v represent the spatial frequency coordinates. If a photographic plate is inserted in plane P_2, the amplitude transmittance of the recorded transparency is given by

$$t = k(|H|^2 + |G|^2) + kHG^*e^{-i4\pi v b} + kH^*Ge^{i4\pi v b}. \qquad (2)$$

where k is a proportionality constant. We note that if a pinhole replaces the input object h, then (2) can be written as

$$t_m = k(A^2 + |G|^2) + kAG^*e^{-i4\pi v b} + kAGe^{i4\pi v b}, \qquad (3)$$

which is the well-known Vander Lugt matched filter. It is apparent that the MF can be generated by a joint transform architecture.

The principle of the correlation operation can be seen by placing a MF in P_2 of an optical processing system shown in Fig. 2. This arrangement is known as a Vander Lugt correlation.

Assuming that the input object is $f(x, y)$, then the complex light distribution behind the correlation filter is given by

$$U_2(u, v) = kF(|H|^2 + |G|^2)$$
$$+ kFHG^*e^{-i4\pi 2 v b} + kFH^*Ge^{i4\pi 2 v b}. \qquad (4)$$

The corresponding output light field at P_3 is

$$U_3(x, y) = U_0$$
$$+ kf(x_3, y_3) * g^*(-x_3, y_3) * h(x_3, y_3) * \delta(x_3, y_3 - 2b)$$
$$+ kf(x_3, y_3) * g(x_3, y_3) * h^*(-x_3, -y_3) * \delta(x_3, y_3 + 2b). \qquad (5)$$

where $*$ denotes the convolution operation, U_0 represents the amplitude distribution centered around the optical axis and the second and third terms are diffracted around $(0, 2b)$ and $(0, -2b)$ respectively. We note that in the filter construction step, we replace $h(x, y)$ by a delta function, i.e., $h(x, y - b) = A\delta(x, y - b)$, the second term of (5) becomes

$$R(x_3, y_3 - 2b)$$
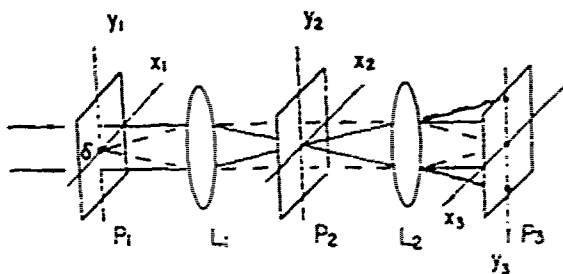$$= \int_{-\infty}^{\infty} \int f(\xi, \zeta) g^*(\xi - x_3, \zeta - y_3 + 2b) d\xi d\zeta. \qquad (6)$$

Fig. 2. A generalized correlation processing system

which is the cross correlation function between f and g. It is apparent that the cross-correlation function R would be diffracted around $(0, 2b)$.

Similarly, for the third term of (5), it becomes

$$R(x_3, y_3 + 2b) = \int_{-\infty}^{\infty} \int f(\xi, \zeta) g(x_3 - \xi, y_3 - \zeta + 2b) d\xi d\zeta. \qquad (7)$$

which is diffracted around $(0, -2b)$.

On the other hand, if the object function $f(x, y)$ in the signal processing step is a δ function, the second and the third terms become

$$R(x_3, y_3 - 2b)$$
$$= h(x_3, y_3) * g^*(-x_3, -y_3) * \delta(x_3, y_3 - 2b).$$
and $\qquad (8)$
$$R(x_3, y_3 + 2b)$$
$$= h^*(-x_3, -y_3) * g(x_3, y_3) * \delta(x_3, y_3 + 2b).$$

for which two cross-correlation functions of h and g are diffracted around $(0, 2b)$ and $(0, -2b)$, respectively. Since the input object is a delta function, as shown in Fig. 2, the correlation filter is illuminated by a collimated light beam. We see that no alignment is required for the correlation filter, and the process is known as a joint transform correlation.

2. Noise Models

Needless to say, the correlation detection would be severely degraded by noise. Noise in a coherent optical correlator generally exists in the input transparency, in the spatial filter, and in the system. We shall consider here, however, the noise in the input plane and in the spatial filter. Since spatial light modulators (SLM) are used as the input transducer and the correlation filter for real-time operation, the noise generated in these devices should be analyzed. For simplicity, the phase noise and amplitude noise will be treated separately in our analyses. Moreover, the phase noise generally comes from the surface variation of the SLM and the amplifier noise is due to the nonuniform transmittance of the device. Thus the noise at the input of an optical correlator may be written as

$$N(x, y) = |N(x, y)| e^{-ip(x, y)}. \qquad (9)$$

where $|N(x, y)|$ and $p(x, y)$ are amplitude and phase noise, respectively.

By referring to the generalized correlator of Fig. 1, we assume that two magneto-optic devices (MODs) are used as input transducers and a liquid crystal light valve (LCLV) is employed as a square law converter in the Fourier plane [9, 10]. Since the MOD is a binary transmission device in which the data written into the

device can be represented by

$$h(x, y) = \sum_{m,n} A_{m,n} |N_{m,n}| e^{i2\pi P_{m,n}}$$

$$\times \operatorname{rect}\left(\frac{x - md}{l}\right) \operatorname{rect}\left(\frac{y - nd}{l}\right). \tag{10}$$

where $A_{m,n}$ is the (0 or 1) binary amplitude transmittance. $|N_{m,n}| e^{i2\pi P_{m,n}}$ is the multiplicative noise introduced by the device. l is the pixel size. d is the period of the discrete structure of the device.

We note that the amplitude noise can be expressed in an optical sense as

$$|N_{m,n}| = \frac{2K}{K+1}\left(\frac{1}{2} - \hat{N}_{m,n}\right), \tag{11}$$

where $\hat{N}_{m,n}$ follows a certain probability distribution. $|\hat{N}_{m,n}| \leq 1/2$. The average value of amplitude noise is $\Delta N = \frac{K}{K+1}$, $K \geq 1$. $\hat{N}_{m,n}$ has random values between $-\frac{1}{2K}$ and $\frac{1}{2K}$. The maximum variation of $|N_{m,n}|$ is from $\frac{K-1}{K+1}$ to 1, which shows the nonuniformity of the transmittance of the device. As an example, if the uniformity of the transmittance of the device is 80%, then the average value of amplitude noise is $\Delta N = 4/5$.

For phase noise, $P_{m,n}$ represents a departure of the thickness of the (m, n) cell from the average thickness of the MOD. We assume that $P_{m,n}$ follows a certain probability distribution and takes on values between $-1/2K$ and $1/2K$, e.g., the random value range is $\Delta P = 1/K$ for $K \geq 1$. For example, if the surface of the device has a flatness of $\lambda/4$, then $\Delta P = 1/4$. The MOD is used as an example for noise model in a real-time correlator. But if the quantity l (or d) represents the minimum resolving cell of an input, the above analysis can be applied to evaluate noise in any optical system.

3. Contrast of Correlation Filter

The intensity distribution of the correlation filter $I(u, v)$ of (1) can be rewritten as

$$I(u, v) = |H|^2 + |G|^2 - 2|H||G| \cos[4\pi b v - (\phi - \varphi)]. \tag{12}$$

where $H(u, v)$ and $G(u, v)$ are the Fourier spectra of $h(x, y)$ and $g(x, y)$, respectively. Then, the contrast of the interference fringe pattern may be represented by

$$\gamma = \frac{2|H||G|}{|H|^2 + |G|^2} \tag{13}$$

It is trivial that if the input object and the reference function are identical i.e., $H(u, v) = G(u, v)$, the contrast of the correlation filter reduces to unity and the irradiance of (12) becomes a sinusoidal grating function. The diffraction efficiency of the grating is of course proportional to its contrast $\gamma (0 \leq \gamma \leq 1)$. Thus, an ideal interference fringe pattern with the highest contrast is required to obtain the best correlation result, i.e., the highest correlation peak. However, in practice, the input object can never be strictly the same as the reference function because of noise disturbances in the devices.

If we further consider the scale change and the rotation, the contrast γ would certainly affect these factors. Assuming the noise exists in the input object, but not in the reference function, then γ should be affected by the amplitude noise ΔN as well as the phase noise ΔP. Since a correlation filter works only in a certain spatial frequency band $(-u_m, u_m), (-v_m, v_m)$, we shall integrate the contrast γ over this band, and the average contrast $\bar{\gamma}$ of the correlation filter over the given frequency band can be expressed as,

$$\bar{\gamma} = \frac{1}{4u_m v_m} \int_{-u_m}^{u_m} \int_{-v_m}^{v_m} \frac{2|H||G|}{|H|^2 + |G|^2} \, du \, dv. \tag{14}$$

4. Noise Effects on Contrast Reduction

Referring to the correlation filter synthesis of Fig. 1, two MODs are used as input transducers. One device is assumed to be contaminated with noise and the another is considered to be without noise. Since the MOD is a binary transmission device, each pixel can be assumed to have a uniform transmittance of amplitude and phase. From (10), the amplitude transmittance $h(x, y)$ with noise and the amplitude transmittance $g(x, y)$ without noise can be expressed in terms of the Fourier transforms. i.e.,

$$H(u, v) = \sum_{m,n} l^2 A_{m,n} |N_{m,n}| e^{i2\pi P_{m,n}}$$

$$\times \operatorname{sinc}(lu)\operatorname{sinc}(lv) e^{-i2\pi(mdu - ndv)} e^{i2\pi bv}. \tag{15}$$

and

$$G(u, v) = \sum_{m,n} l^2 A_{m,n} \operatorname{sinc}(lu)\operatorname{sinc}(lv) e^{-i2\pi(mdu - ndv)} e^{-i2\pi bv}. \tag{16}$$

where $N_{m,n} e^{i2\pi P_{m,n}}$ represents the noise. $A_{m,n}$ has the binary values 1 or 0, and $2b$ denotes the separation of the two devices in the input plane. For simplicity, only the 1-D case is considered. From (1), the intensity distribution along the v axis ($u = 0$) in the recording

plane can be expressed by

$$I = I^4 \mathrm{sinc}^2(lv)\left[\left(\sum_n A_n \cos 2\pi n dv\right)^2 \right.$$

$$\div \left(\sum_n A_n \cos 2\pi n dv\right)^2$$

$$+ \left(\sum_n A_n |N_n| \cos 2\pi (P_n - ndv)\right)^2$$

$$\div \left(\sum_n A_n |N_n| \sin 2\pi (P_n - ndv)\right)^2$$

$$\left. \div 2 \sum_{n,n'} A_n |N_n| A'_n \cos 2\pi [P_n - (n-n')dv \div 2bv]\right].$$

$$(17)$$

Let us first consider the amplitude noise, i.e., $P_n = 0$, for all $n = \pm 1, \pm 2, \dots$. Assuming that $A_n = 1$ for all $n = \pm 1, \pm 2, \dots$, equation (17) can be reduced to

$$I = I^4 \mathrm{sinc}(lv)^2 \left[\left(\sum_n \cos 2\pi n dv\right)^2 + \left(\sum_n |N_n| \cos 2\pi n dv\right)^2 \right.$$

$$\div \left(\sum_n |N_n| \sin 2\pi n dv\right)^2$$

$$\left. \div 2 \sum_n \cos 2\pi n dv \sum_n |N_n| \cos 2\pi (ndv - 2bv)\right]. \qquad (18)$$

If $|N_n| = 1$, the intensity of (18) is reduced to

$$I = 2I^4 \mathrm{sinc}^2(lv)\left(\sum_n \cos 2\pi n dv\right)^2 (1 \div \cos 4\pi bv). \qquad (19)$$

and the contrast of the filter (without noise) for the above equation is equal to 1. Equation (18) is used to calculate the intensity distribution of the interference fringes of a correlation filter, without amplitude noise $|N_n| = 1$, and with noise, $|N_n| = \dfrac{2K}{K+1}\left(\dfrac{1}{2} - \hat{N}_n\right)$, where $K \geq 1$. It is assumed that $\hat{N}_n$ is a random variable taken from a range between $-1/2K$ and $1/2K$. We further assume that $\hat{N}_n$ obeys one of the probability distributions such as Gaussian, uniform or exponential.

If the noise introduced by the device is only phase noise, i.e., $|N_n| = 1$ for all $n = \pm 1, \pm 2, \dots$, and assuming that $A_n = 1$ for all n, then (18) becomes

$$I = I^4 \mathrm{sinc}^2(lv)\left[\left(\sum_n \cos 2\pi n dv\right)^2 \right.$$

$$\div \left(\sum_n \cos 2\pi (P_n - ndv)\right)^2 \div \left(\sum_n \sin 2\pi (P_n - ndv)\right)^2$$

$$\left. \div 2 \sum_n \cos 2\pi n dv \sum_n \cos 2\pi (P_n - ndv \div 2bv)\right]. \qquad (20)$$

where P_n is a random variable taken from a range between $-1/2K$ and $1/2K$, which may also obey some probability distribution.

Three typical noise statistics are considered: Gaussian, uniform, and exponential. In each case, the

correlation peak intensity is compared with the case without phase noise.

As an example, assuming that N_n and P_n follow a truncated Gaussian density distribution,

$$\mathrm{rect}\left[\frac{x-a}{4\sigma}\right]\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}. \qquad (21)$$

It is assumed that $a = 0$ and that if $\hat{N}_n$ and P_n take values between $-1/2K$ and $1/2K$, then we have

$$2\sigma = \frac{1}{2K}, \qquad K \geq 1. \qquad (22)$$

For simplicity, the input and reference functions are 1-D arrays of dimension $2L$. Since the spectral distribution of a rectangular function consists of infinite lobes of finite bandwidth $1/2L$, it can be treated as a periodic function in the spatial frequency domain.

It can be seen that the contrast of the filter is reduced by the amplitude noise, while for the phase noise, contrast is reduced and the fringes of the filter are shifted. If the average noise ΔN decreases from $1/2$ to $1/2K$, where $K \geq 1$, and $\hat{N}_n$ obeys a probability distribution, such as Gaussian, uniform, or exponential, the average contrast is a function of the average amplitude noise as shown in Fig. 3. The curve showing contrast reduction due to Gaussian amplitude noise seems to be a reasonable description of amplitude noise introduced by nonuniform transmittance of the device. When the average value $\Delta N \leq 0.24$ (i.e., uniform transmittance equals 76%), the contrast of the correlation filter drops 10% from unity. Figure 4 shows the contrast as a function of the random value range ΔP for various types of statistical phase noise (e.g., Gaussian, uniform, and exponential distribu-
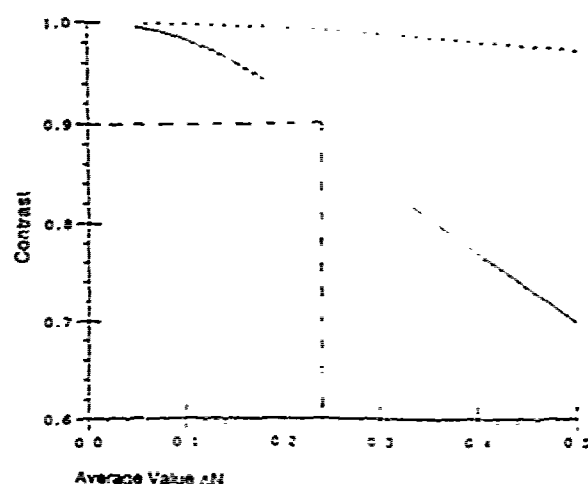

Fig. 3. Average contrast of the correlation filter as a function of the average amplitude noise ΔN for Gaussian (——), uniform (- - -), and exponential (·····) amplitude noise models
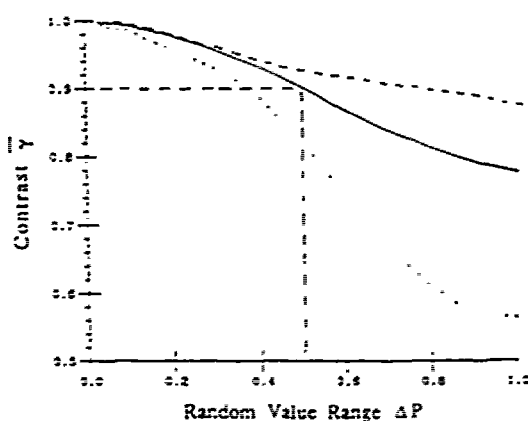
Fig. 4. Average contrast of the correlation filter as a function of the random value range ΔP, for Gaussian (——), uniform (·····), and exponential (---) phase noise models

Assuming that the scale of the input function is changed, as can be expressed by

$$h(x, y) = \text{rect}\left(\frac{Sx}{2L_x}\right)\text{rect}\left(\frac{S(y+b)}{2L_y}\right) H_0(u_0, v_0)e^{i2\pi(u_0x+v_0y)},$$

(25)

where S is the scale-change factor, and H_0 is a weighting factor, then the corresponding distribution at the Fourier plane is given by

$$H(u, v) = \frac{4L_xL_y}{S^2} H_0(u_0, v_0)\text{sinc}\left[2L_x\left(\frac{u}{S}-u_0\right)\right]$$

$$\times \text{sinc}\left[2L_y\left(\frac{v}{S}-v_0\right)\right]e^{i2\pi v}.$$

(26)

By referring to (13), the contrast of the correlation filter (assuming that $H_0 = G_0$), can be written as

$$\gamma = \frac{2\,\text{sinc}[2L_x(u-u_0)]\,\text{sinc}\left[2L_x\left(\frac{u}{S}-u_0\right)\right]\text{sinc}[2L_y(v-v_0)]\,\text{sinc}\left[2L_y\left(\frac{v}{S}-v_0\right)\right]}{S^2\,\text{sinc}^2[2L_x(u-u_0)]\,\text{sinc}^2[2L_y(v-v_0)] + \frac{1}{S^2}\text{sinc}^2\left[2L_x\left(\frac{u}{S}-u_0\right)\right]\text{sinc}^2\left[2L_y\left(\frac{v}{S}-v_0\right)\right]}.$$

(27)

tions). If $\Delta P = 1/2$, i.e., the surface roughness of the devices is equal to or smaller than $\lambda/2$, the contrast of the filter is about 0.9, for Gaussian noise. From Figs. 3 and 4, we have seen that a contrast of 0.9 yields $\Delta N = 0.24$ for amplitude noise and $\Delta P = 1.2$, for phase noise. This means that the surface quality of $\lambda/2$ causes the contrast reduction as a transmittance uniformity of 76%.

5. Effect Due to Differences Between Object and Reference Functions

To consider the contrast reduction of the correlation filter caused by such differences, the reference function distributed in the Fourier plane is given by

$$g(x, y) = \sum_{m,n} G_{m,n}e^{i2\pi(m\frac{l}{L_x}x + n\frac{l}{L_y}y)}.$$

(23)

where $2L_x \times 2L_y$ are the dimensions of the input functions. $G_{m,n}$ is the Fourier coefficient, and l is the pixel size of the input function. However, for simplicity, a pair of spatial frequencies (u_0, v_0) is considered, which is given by

$$g_r(x, y) = G(u_0, v_0)e^{i2\pi(u_0x+v_0y)}\text{rect}\left(\frac{x}{2L_x}\right)\text{rect}\left(\frac{y}{2L_y}\right).$$

(24)

where $u_0 = m_0\frac{l}{L_x}$ and $v_0 = n_0\frac{l}{L_y}$.

To obtain a simple and clear physical picture, a 1-D case where $u=0$ and $u_0=0$ is considered. Thus (27) reduces to

$$\gamma = \frac{2\,\text{sinc}(2L_yv)\,\text{sinc}\left(2L_y\frac{v}{S}\right)}{S^2\,\text{sinc}^2(2L_yv) + \frac{1}{S^2}\text{sinc}^2\left(2L_y\frac{v}{S}\right)}.$$

(28)

The average contrast $\bar{\gamma}$ can of course be obtained by using (14). Figure 5 shows the average contrast $\bar{\gamma}$ as a function of scale factor S over the frequency band
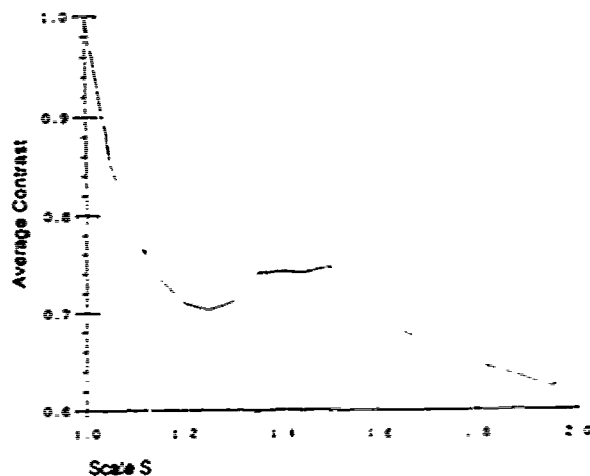

Fig. 5. Average contrast as a function of scale change S over a frequency bandwidth (4 periods)

$(0, r_m)$. We see that the average contrast of the filter decreases as the scale factor increases.

On the other hand, if the input function is rotated as described by the following equation:

$$i(x, y) = \text{rect}\left(\frac{x\cos\theta + (y + b)\sin\theta}{2L_x}\right)$$

$$\times \text{rect}\left(\frac{-x\sin\theta + (y + b)\cos\theta}{2L_y}\right)$$

$$\times H_0(u_0, r_0)e^{i2\pi u_0(x\cos\theta + y\sin\theta)}e^{i2\pi r_0(-x\sin\theta + y\cos\theta)}. \quad (29)$$

where r is the angle of rotation of the input function. The corresponding Fourier spectral distribution is

$$H(u, r) = 4L_x L_y H_0(u_0, r_0)\,\text{sinc}\,2L_x(q\sin\theta$$

$$+ p\cos\theta)\sin 2L_y(q\cos\theta - p\sin\theta)e^{i2\pi b r}. \quad (30)$$

where $p = r - u_0\sin\theta - r_0\cos\theta$ and $q = u - u_0\cos\theta + r_0\sin\theta$. Then the average contrast of the filter, for a 1-D representation is given by

$$\gamma = \frac{2\,\text{sinc}(2L_y r)\,\text{sinc}(2L_x r\cos\theta)\,\text{sinc}(2L_y r\sin\theta)}{\text{sinc}^2 2L_y r + \text{sinc}^2(2L_x r\cos\theta)\,\text{sinc}^2(2L_y\sin\theta)}. \quad (31)$$

where it is assumed that $r_0 = 0$.

Three average contrasts for different frequency bands (4, 8, and 16 periods) are plotted in Fig. 6 as a function of rotational angle θ. We see that the larger the frequency band, the faster the average contrast decreases.
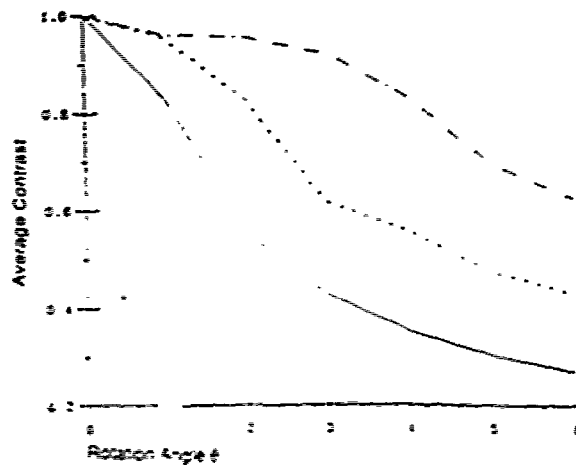
6. Contrast Reduction on MF

It is evident that a matched filter can be constructed by replacing the input function with a pinhole, i.e., $h(x, y - b) = A\delta(x, y - b)$ as shown in Fig. 1.

The contrast of MF can be written as

$$\gamma = \frac{2A|G|}{A^2 + |G|^2}. \quad (32)$$

where A is the constant amplitude of the tilted reference plane wave. Thus for a MOD, the contrast of the MF can be written as

$$\gamma = \frac{2A4L_x L_y|G_0(u_0, r_0)\,\text{sinc}[2L_x(u - u_0)]\,\text{sinc}[2L_y(r - r_0)]|}{A^2 + |4L_x L_y G_0(u_0, r_0)\,\text{sinc}[2L_x(u - u_0)]\,\text{sinc}[2L_y(r - r_0)]|^2}. \quad (33)$$

Again for a 1-D representation, we have

$$\gamma = \frac{2R|\text{sinc}[2L_y(r - r_0)]|}{R^2 + \text{sinc}^2[2L_y(r - r_0)]}. \quad (34)$$

where $R = \dfrac{A}{4L_x L_y G_0}$. From this equation we see that the contrast is a function of the spatial frequency variable r, where r_0 is the carrier spatial frequency of the reference plane wave. Thus the contrast, over a given frequency band $(-r_m + r_m)$, can be expressed as

$$\bar{\gamma} = \frac{1}{2r_m}\int_{-r_m}^{r_m}\frac{2R|\text{sinc}[2L_y(r - r_0)]|}{R^2 + \text{sinc}^2[2L_y(r - r_0)]}\,dr. \quad (35)$$

To obtain an optimum $\bar{\gamma}$ over a given frequency bandwidth $(0, r_m)$, we shall search for the best R. Figure 7 shows the $\bar{\gamma}$ as a function of frequency periods over the range $(0, r_m)$. Figure 8 is the plot of $\bar{\gamma}$ (over 8
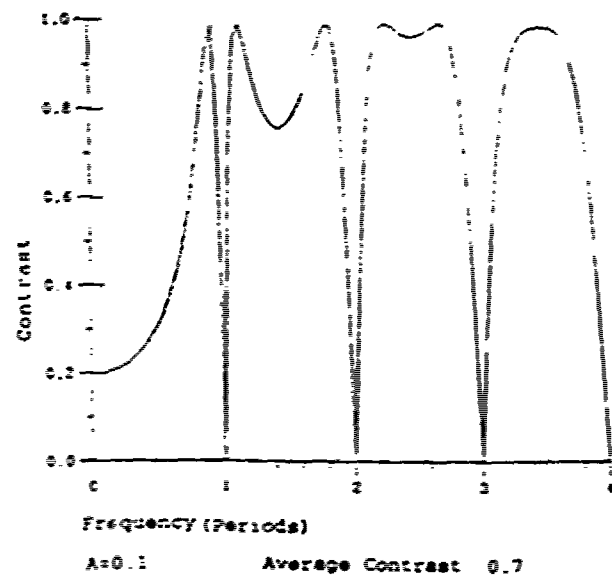

Fig. 6. Average contrast over different frequency bandwidths (4, ---, 8, ····, 16, ——, periods) as a function of rotation angle θ

A=0.1   Average Contrast 0.7

Fig. 7. Average contrast of a matched filter as a function of frequency for object-to-reference beam ratio equal to 1:3
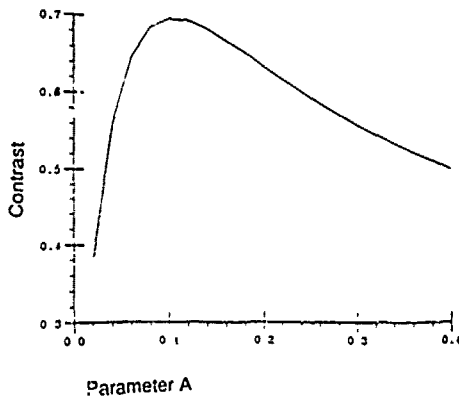
Fig. 8. Average contrast of a matched filter (over 8 periods) as a function of the intensity ratio $R = A/(4L_x L_y G_0)$

Table 1. Optimum contrast of a matched filter

No. of periods	2	4	8	12	16	20
Optimum $\bar{\gamma}$	0.69	0.7	0.71	0.71	0.71	0.71
Reference-to-object ratio	69	12	3	1.9	0.75	0.5

periods) as a function of A, which is proportional to R. From this figure, we see that an optimum $\bar{\gamma} = 0.71$ occurred at $R = 0.05$. As an example, when $4L_x L_y = 7.2 \times 4.8\,\mathrm{mm}^2$, then $\frac{A}{G} \cong 1.7$. That means the intensity reference-to-object beam ratio is about 3. We note that over a different number of periods, the optimum $\bar{\gamma}$ changes very slightly with R.

The corresponding values of optimum $\bar{\gamma}$ reference-to-object beam ratio, and the number of periods is tabulated in Table 1 for $4L_x L_y = 7.2 \times 4.8\,\mathrm{mm}^2$. From this table we see that the average contrast of the matched filter is around 0.7 for different numbers of periods covered by the MF. Comparing with the noisy JTC filter of Figs 3 and 4, we see that the contrast for the JTC filters are higher than the noiseless MF, if the surface quality of the device is better than λ, and the uniformity of the device is higher than 50%.

7. Correlation Degradation

It is well known, both for the Vander Lugt correlator and the JTC, that the resulting output correlation appears as a bright spot. To determine the performance of an optical correlation, it is necessary to measure the correlation peak intensity.

Let the correlation functions of $f(x, y)$ and $g(x, y)$ be given by

$$R(x, y) = \int \int_{\infty} F(u, v)\, G^*(u, v)\, e^{i2\pi(ux + vy)}\, du\, dv. \tag{36}$$

It can be shown that the following well-known inequality holds:

$$|R(0, 0)|^2 \geqq |R(x, y)|^2, \tag{37}$$

where $|R(0, 0)|^2$ is defined as the correlation peak intensity.

As before, two MODs are used as reference and input devices in the correlator to analyze the degradation of the output peak intensity. Assuming that the input function $f(x, y)$ is a scale variant of the reference function $g(x, y)$, we have

$$f_s(x, y) = \sum_{m, n} A_{m, n} \operatorname{rect}\left(\frac{x - md}{l/S}\right) \operatorname{rect}\left(\frac{y - nd}{l/S}\right), \tag{38}$$

and the corresponding Fourier transform is

$$F_s(u, v)$$
$$= \sum_{m, n} \left(\frac{l}{S}\right)^2 A_{m, n} \operatorname{sinc}\left(\frac{l}{S} u\right) \operatorname{sinc}\left(\frac{l}{S} v\right) e^{-i2\pi(mdu + ndv)}, \tag{39}$$

where S is the scale-change factor. Thus the correlation function of $f(x, y)$ and $g(x, y)$ can be written as

$$R_s(x, y) = \int \int_{-\infty}^{\infty} \left(\frac{l^2}{S}\right)^2 \operatorname{sinc}(lu) \operatorname{sinc}(iv)$$
$$\times \operatorname{sinc}\left(\frac{l}{S} u\right) \operatorname{sinc}\left(\frac{l}{S} v\right)$$
$$\times \sum_{m, n} A_{m, n} e^{-i2\pi d(mu + nv)}$$
$$\times \sum_{m, n} A_{m, n} e^{i2\pi \frac{d}{S}(mu + nv)} e^{i2\pi(ux + vy)} du\, dv. \tag{40}$$

Using the convolution theorem, (40) can be expressed as

$$R_s(x, y) = r_s(x)\, r_s(y) * \sum_{m, n} \sum_{m', n'} A_{m, n} A_{m', n'}$$
$$\times \delta\left[x + d\left(\frac{m'}{S} - m\right)\right] \delta\left[y + d\left(\frac{n'}{S} - n\right)\right], \tag{41}$$

where

$$r_s(x) = \operatorname{rect}\left(\frac{x}{l}\right) * \operatorname{rect}\left(\frac{x}{l/S}\right)$$

$$= \begin{cases} \dfrac{l}{S}, & |x| \leqq \dfrac{l}{2}\left(1 - \dfrac{1}{S}\right) \\[2mm] \dfrac{l}{S} - \dfrac{2}{1+S}|x|, & \dfrac{l}{2}\left(1 - \dfrac{1}{S}\right) < |x| < \dfrac{l}{2}\left(1 + \dfrac{1}{S}\right) \\[2mm] 0, & |x| \geqq \dfrac{l}{2}\left(1 + \dfrac{1}{S}\right) \end{cases}.$$

$$\tag{42}$$

For example, a letter "A" is displayed on a MOD device as the reference function represented by $\sum_{m,n} A_{m,n}$. The correlation peak intensity $R_s^2(0,0)$ of "A" and the scale-changed "A" are calculated as a function of the scale change factor S. Again using a 1-D presentation and assuming that $A_{m,n}=1$ for all m and n, (40) can be reduced to

$$R_s(x)= \int_{-\infty}^{\infty} \frac{l^2}{S}\operatorname{sinc}(lu)\operatorname{sinc}\left(\frac{l}{S}u\right)$$

$$\times \sum_m e^{-i2\pi dmu}\sum_{m'}e^{i2\pi\frac{d}{S}m'u}e^{i2\pi ux}du$$

$$=r_s(x)*\sum_{m,m'}\delta\left[x+d\left(\frac{m'}{S}-m\right)\right] \qquad (43)$$

Using this result, the correlation peak intensity can be computed. Figure 9 shows the correlation peak intensity, $R_s^2(0,0)$, as a function of the scale-changed "A" as the dotted curve, while the solid curve represents $R_s^2(0)$ for an 1-D representation.

If the input function $f(x,y)$ is rotated as given by

$$f_r(x,y)= \sum_{m,n} A_{m,n}\operatorname{rect}\left(\frac{x\cos\theta+y\sin\theta-md}{l}\right)$$

$$\times \operatorname{rect}\left(\frac{-x\sin\theta+y\cos\theta-nd}{l}\right), \qquad (44)$$

where θ denotes the angle of rotation with respect to the reference function. The corresponding spectral distribution can be written as

$$F_r(u,v)= \sum_{m,n} l^2 A_{m,n}\operatorname{sinc}[l(u\cos\theta+v\sin\theta)]$$

$$\times \operatorname{sinc}[l(-u\sin\theta+v\cos\theta)]$$

$$\times e^{-i2\pi md(u\cos\theta+v\sin\theta)}e^{-i2\pi nd(-u\sin\theta+v\cos\theta)} \qquad (45)$$

Scale S

48x48 Array— . "A" on 11x15 Array---
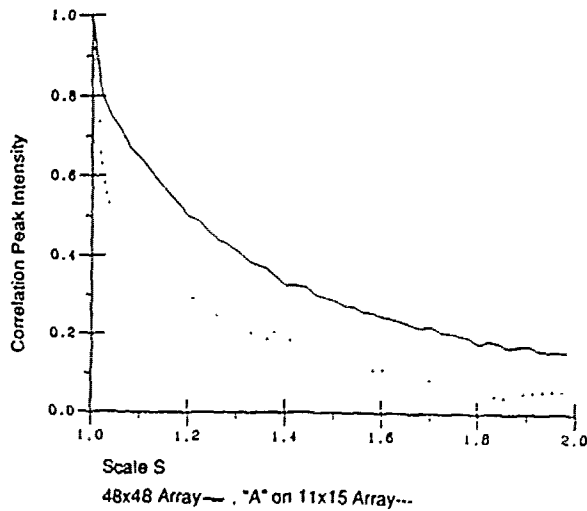
Fig. 9. Correlation peak intensity as a function of the scale change factor S
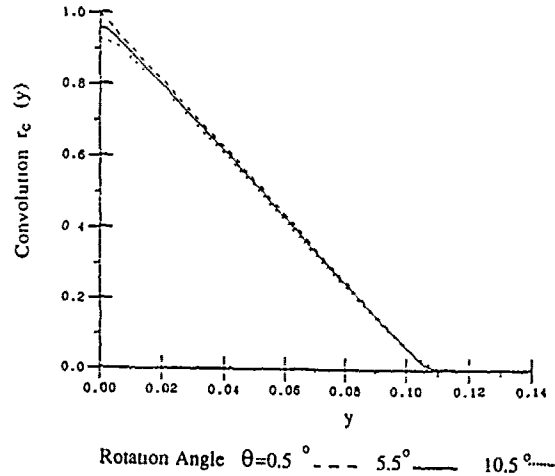

Fig. 10. Convolution of a square with $d=0.127$ mm and its rotated version

For simplicity, consider the correlation of the reference function and its rotated version only on the y axis (i.e., $u=0$); then the correlation function can be expressed as

$$R_r(y)=r_c(y)*\sum_{m,n}\sum_{m',n'}A_{m,n}A_{m',n'}$$

$$\times \delta[y+d(m\sin\theta+n\cos\theta-n')], \qquad (46)$$

where $r_c(y)$ represents the convolution between a square (with side l) with respect to its rotated version along the y axis, i.e.,

$$r_c(y)=\operatorname{rect}\left(\frac{x}{l}\right)\operatorname{rect}\left(\frac{y}{l}\right)*\operatorname{rect}\left(\frac{x\cos\theta+y\sin\theta}{l}\right)$$

$$\times \operatorname{rect}\left(\frac{-x\sin\theta+y\cos\theta}{l}\right). \qquad (47)$$

Figure 10 sketches the function $r_c(y)$ for different rotation angles, $0.5°$, $5.5°$, and $10.5°$. From this figure we see that for small θ, $r_c(y)$ can be approximated by
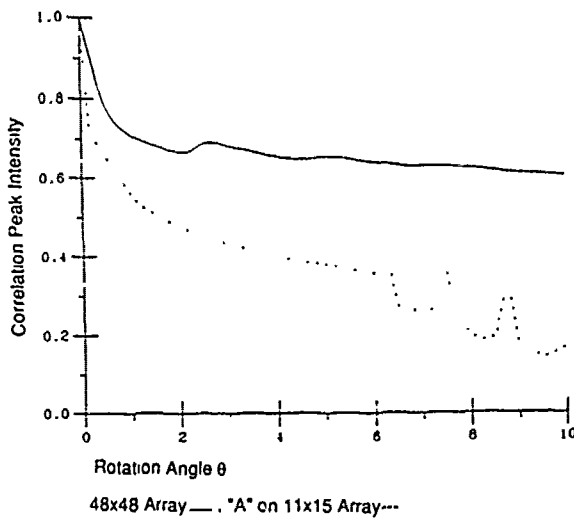
$$r_c(y)\approx \wedge\left(\frac{y}{l}\right)\doteq \begin{cases} l-|y|, & |y|<l, \\ 0, & |y|\geqq l. \end{cases} \qquad (48)$$

Again, we use the letter "A" for computing the correlation intensity $R_r^2(0)$. Figure 11 shows the correlation peak intensity as a function of rotation angle θ (dotted curve).

If it is assumed that $A_{m,n}=1$, then (46) reduces to

$$R_r(y)=r_c(y)*\sum_{m,n,n'}\delta[y+d(m\sin\theta+n\cos\theta-n')], \qquad (49)$$

which is essentially the correlation of the square array and its rotated function. The correlation peak intensity $R_r(0)$ is plotted as the solid curves in Fig. 11.

Fig. 11. Correlation peak intensity vs. rotation angle θ

48x48 Array___ . "A" on 11x15 Array---

Fig. 12. Correlation peak intensity as a function of average value ΔN for Gaussian (——), uniform (····), and exponential (---) amplitude noise

8. Noise and Correlation Degradation

By referring to (10), the output correlation of the noisy input object with the reference function is given by

$$R_N(x,y) = \sum_{m,n}\sum_{m',n'} A_{m,n}A_{m',n'}|N_{m,n}|e^{i2\pi P_{m,n}}$$
$$\times \wedge\left[\frac{x-d(m-m')}{l}\right]\wedge\left[\frac{y-d(n-n')}{l}\right],$$

(50)

where $\wedge$ is defined in (48).

If only amplitude noise is considered with $A_{m,n}=A_{m',n'}=1$, the above equation is reduced to

$$R_a(x,y) = \sum_{m,n}\sum_{m',n'} |N_{m,n}| \wedge\left[\frac{x-d(m-m')}{l}\right]$$
$$\times \wedge\left[\frac{y-d(n-n')}{l}\right],$$

(51)

where $|N_{m,n}|$ is given by (11).

Similarly, for phase noise only, the output correlation function becomes

$$R_p(x,y) = \sum_{m,n}\sum_{m',n} e^{i2\pi P_{m,n}} \wedge\left[\frac{x-d(m-m')}{l}\right]$$
$$\times \wedge\left[\frac{y-d(n-n')}{l}\right],$$

(52)

where $e^{i2\pi P_{m,n}}$ represents the phase noise. Figure 12 shows a set of correlation peak intensities, $R_a^2(0,0)$, as a function of the average value ΔN, for Gaussian, uniform, and exponential amplitude noise statistics, respectively From this figure, we see that the variation of peak intensities for those noise statistics, drop rather rapidly at about the same rate. The peaks drop to about 77% when $\Delta N = 0.12$. The average value of 0.12

Fig. 13. Correlation peak intensity as a function of the random value range ΔP for Gaussian (——), uniform (····), and exponential (---) phase noise

means the maximum departure of the amplitude transmittance of a MOD pixel. This corresponds to about 24% of the unity amplitude transmittance, which is equal to 0.19 optical density.

Similarly, $R_p^2(0,0)$ as a function of the random value range ΔP, for various phase noise statistics is plotted in Fig. 13. It is seen that the Gaussian phase noise is a reasonable description. When the value range is $\Delta P = 1/2$, i.e., the device has a surface quality of $\lambda/2$, the correlation peak intensity drops 23% from the unity correlation peak. Comparing Figs. 12 and 13, we see that the phase noise due to a surface quality of $\lambda/2$ has the same effect on the correlation peak intensity as an amplitude noise with the average value of 0.12. This corresponds to a transmittance uniformity of 88% of the MOD.

9. Noise Performance

Assume that R and R_N are the output correlation functions with and without noise respectively. The output signal-to-noise ratio can be defined as

$$\frac{S}{N} = \frac{E\{R^2\}}{E\{(R-R_N)^2\}}, \tag{53}$$

where $E\{\ \}$ represents the ensemble average. However, in an optical sense, the signal-to-noise ratio can be written as

$$\frac{S}{N} = \frac{\iint\limits_{XY} |R|^2 dx\, dy}{\iint\limits_{XY} |R-R_N|^2 dx\, dy}, \tag{54}$$

where (x, y) represent the output coordinate system.

Again for a 1-D representation, R and R_N, as referred to in (53), are given by

$$R_N = \sum_{n,n'} |N'_n| e^{i2\pi P_n} \wedge \left[\frac{y-d(n-n')}{l}\right], \tag{55}$$

and

$$R = \sum_{n,n'} \wedge \left[\frac{y-d(n-n')}{l}\right]. \tag{56}$$

Thus, we have

$$R - R_N = \sum_{n,n'} \wedge \left[\frac{y-d(n-n')}{l}\right](1 - |N_n|e^{i2\pi P_n}). \tag{57}$$

If the amplitude noise is considered, then we have

$$R - R_N = \sum_{n,n'} \Delta N(1 + 2K\hat{N}_n) \wedge \left[\frac{y-d(n-n')}{l}\right]. \tag{58}$$

Similarly for phase noise,

$$|R - R_N|^2 = \left(\sum_{n,n'} \wedge \left[\frac{y-d(n-n')}{l}\right]\right.$$
$$\left. - \sum_{n,n'} \cos 2\pi P_n \wedge \left[\frac{y-d(n-n')}{l}\right]\right)^2$$
$$+ \left(\sum_{n,n} \sin 2\pi P_n \wedge \left[\frac{y-d(n-n')}{l}\right]\right)^2. \tag{59}$$

By assuming that $\bar{N}_n$ and P_n are Gaussianly disturbed, the output signal-to-noise ratios for amplitude and phase noise disturbances are plotted in Figs. 14 and 15, respectively. For example, if $\Delta P \le 1/4$ (the surface quality of the device is about $\lambda/4$ or better), the signal-to-noise ratio should be greater than 23 dB. For the same SNR in Fig. 14, $\Delta N \ge 0.07$, which is equivalent to a maximum departure of 14% in the amplitude transmittance of the device, or a transmittance uniformity of 93%.
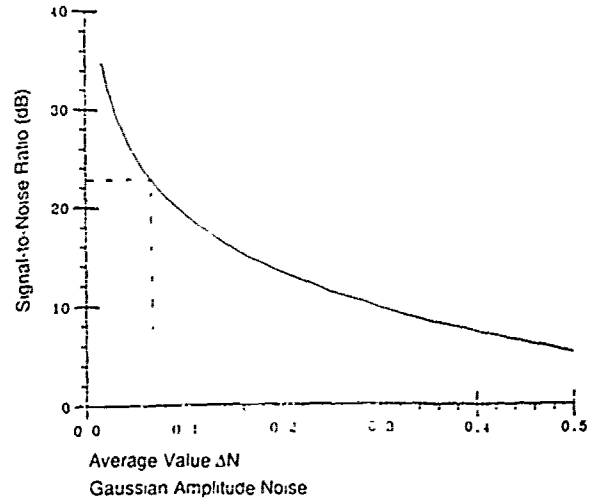

Fig. 14. Signal-to-noise ratio S/N as a function of the average value ΔN for Gaussian amplitude noise
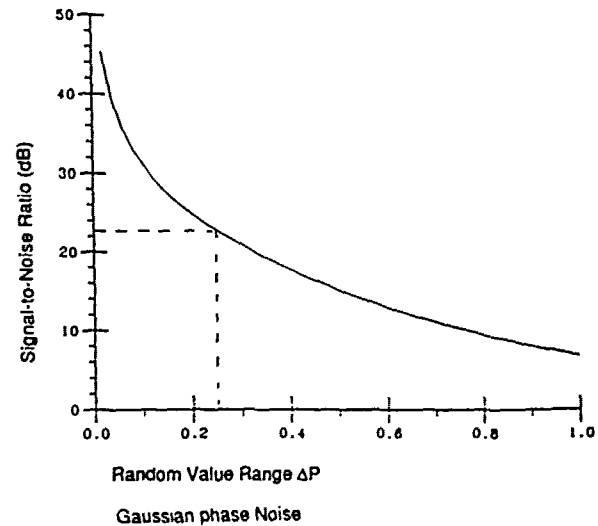

Fig. 15. Signal-to-noise ratio S/N as a function of the random value ΔP due to Gaussian phase noise

10. Comparison of VLC and JTC

Let us now look at the modulation indices in the joint transform filter and the matched filter. If the input function is identical to the reference function under the same intensity illumination, the joint transform interference pattern is fully modulated. In other words, the contrast of the fringe pattern is unity throughout the joint transform filter. However, to estimate the modulation of the Vander Lugt type matched filter, the ensemble average contrast is utilized. The average contrast of the fringes was found to be about 0.7 for an optimum matched filter (i.e., object-to-reference beam ratio of unity) and since the average modulation index

of the joint transform hologram is higher than the matched filter. the joint transform correlator should provide a higher autocorrelation peak intensity.

For the Vander Lugt type. it is required that the matched filter be placed back in its original recording position. Practical alignment must be to within the order of about 0.01 mm in the Fourier plane. However, for the joint transform correlation, there is no alignment requirement for the position of the filter in Fourier plane.

The space × bandwidth product is an important specification for an optical processing system. although it never reaches the specific limit in practice. At first glance. the joint transform correlator requires a space × bandwidth product twice that of the Vander Lugt correlator. where in Fig. 1 a pinhole for the construction of the matched filter can be seen. Therefore. the space × bandwidth product of the Vander Lugt is indeed lower than that of the joint transform correlator. The key issue is. however. not the space × bandwidth product. but that one can fully utilize the product provided by the correlator.

There is also a difference between the Vander Lugt and joint transform correlators in multiplexing. or multiple object recognition. In a Vander Lugt correlator, several functions are placed in the input plane and a set of metched filters placed in the Fourier plane for the multiplexing operation. However, in a joint transform correlator. instead of using multiple filters in the Fourier plane. multiple reference functions can be put in the input plane of the correlator for multiple object recognition. In addition. a multichannel joint transform correlator can be constructed by using a special multiple Fourier transform lens or an array of small lenses

As mentioned previously. the noise in most optical correlation systems is basically of multiplicative type. and is introduced either by the input object or by spatial light modulators used in the optical system. We have shown that the noise effects on the output correlation peak for both Vander Lugt and joint transform correlators are about the same.

We should emphasize that. for real-time operation. the joint transform correlator is simpler and more straightforward to operate than the Vander Lugt correlator. A real-time joint transform correlator is a one-step operation. if the appropriate spatial light modulators are introduced in the input and the Fourier planes. The advantage of the one-step operation is that prefabrication of the correlation filter is not required. However. for the Vander Lugt correlator. a two-step process is required in which a matched filter must be prefabricated. Notice that a computer-generated matched filter can be directly implemented in a spatial light modulator. such as a MOD. Although

Table 2. Comparison of Vander Lugt and joint transform correlators

	Joint transform	Vander Lugt
Modulation index	$\gamma = 1$	$\gamma = 0.7$
Alignment	No	crucial. $\Delta r = 0.01$ mm
Space × bandwidth	Greater	smaller
Multiplexing	Set of reference functions	Set of matched filters
Real time operation	One-step	Two-step
Noise performance	The same	The same

this makes a one-step Vander Lugt correlator possible. complicated computations must be performed to generate the filter.

We summarize the comparison of the Vander Lugt and the joint transform correlator in Table 2. Thus. we see that the joint transform correlator would. in general. perform better. particularly for real-time implementation.

11. Conclusion

A generalized optical correlator has been used for the Vander Lugt and the joint transform correlators. Under such a unified description. the analyses of to the noise disturbances and the differences between the input and the reference functions are valid for both architectures. To consider the effects due to noise. various amplitude and phase models were used. Magneto-optic spatial light modulators were used for the system performance calculation.

The contrast of the correlation filters for both correlators were analyzed. We have shown that the contrast of the correlation filter was reduced by the presence of noise. The effect of scale change and rotation of the input function on the contrast of the filter was also calculated. The correlation peak intensity under the influence of noise and differences between the input and reference function was also studied. We have found that the decrease of the correlation peak depends on the type of noise and the value range of the noise. The scale change and rotation of the input also degrade the correlation peak intensity. By comparing the performances of the Vander Lugt and joint transform correlators. we have shown that the joint transform correlator is easier to construct and offers advantages in real-time implementations.

References

1. A. Vander Lugt: IEEE Trans. IT-10, 139 (1964)
2. C.S. Weaver, J.W. Goodman: Appl. Opt. 5, 1248 (1966)
3. J.E. Rau: J. Opt. Soc. Am. 56, 1490 (1966)
4. A. Vander Lugt: Appl. Opt. 5, 1760 (1966)
5. J. LaMachia, D. White: Appl. Opt. 7, 91 (1968)
6. N. Shi: Opt. Lett. 3, 85 (1978)
7. A. Grumet: *Automatic Target Recognition System.* U.S. Patent 3, 337, 492 (1972)
8. F.T.S. Yu, X.J. Lu: Appl. Opt. 23, 3109 (1984)
9. F.T.S. Yu, X.J. Lu: Opt. Commun. 52, 10 (1984)
10. F.T.S. Yu, J.E. Ludman: Opt. Lett. 11, 395 (1986)
11. D. Casasent: Proc. IEEE 65, 143 (1977)
12. A.D. Gara: Appl. Opt. 16, 149 (1977)
13. B.D. Guenther, C.R. Christensen, J. Upatnieks: IEEE J. QE-15, 1348 (1979)

APPENDIX 10.6

LCTV Color Encoder

Original article

ELECTRICAL ENGINEERING DEPARTMENT. THE PENNSYLVANIA STATE UNIVERSITY

University Park. Pennsylvania 16802

GRAY LEVEL PSEUDOCOLOR ENCODING USING A LIQUID CRYSTAL TELEVISION

F.T.S. YU. S. JUTAMULIA and E. TAM

Codage en fausses couleurs des niveaux de gris par un système de télévision à cristal liquide

SUMMARY : A real-time gray level intensity pseudocolor encoding method utilizing a liquid crystal television (LCTV) is presented. The LCTV is used under two different polarizer/analyser combinations to generate a positive image and a negative image of an object. The images are also encoded with two different primary colors by positioning the color filters at the two optical paths respectively. The superposition of these images produces a gray level pseudocolor encoded image. Discussions on the properties of the LCTV is provided. Preliminary experimental results are also demonstrated.

RESUME : On présente une methode de codage en fausses couleurs de niveaux de gris utilisant en temps reel un systeme de télévision à cristal liquide. Le systeme à cristal liquide est utilisé avec deux combinaisons analyseur-polariseur pour donner de l'objet une image positive et une image negative. Les images sont codées avec deux couleurs de base en plaçant des filtres sur les deux trajets optiques. La superposition de ces images donne un codage des gris en fausses couleurs. Les proprietes du système à cristal liquide sont discutees et on donne quelques resultats experimentaux preliminaires.

1. — INTRODUCTION

Most of the optical images obtained in various scientific applications are gray level intensity images. for example. the scanning electron microscopic images. multispectral band aerial photographic images. X-ray transparencies. endoscopic images. etc. However. human perceives variations in color better than those in gray levels. In other words. a color-coded image provides better visual discrimination [1].

Density pseudocolor encoding by half tone screens implementation with a coherent optical processor was first reported by Liu and Goodman [2]. and later with a white-light processor by Tai *et al.* [3]. Based on similar principles. many optical architectures have been reported recently [4-8].

The recent appearance of a low-cost (≈ $ 100)black-and-white liquid crystal television (LCTV) has attracted a great deal of attention to explore its usefulness for real-time optical signal processing [9] We demonstrated the application of

LCTV to white-light processing [10]. which included the optical pseudocolor encoding. i.e. different grating structures resulted in different colors. A direct method to pseudocolor-encode light intensity using the LCTV was also reported recently [11]. In this method. a light intensity distribution is recorded by a video camera. in which light intensity is transduced to electric signal. The electronic video signal from the camera is then fed to the LCTV to produce the voltage difference across the liquid crystal. A color pattern corresponding to the applied voltage will appear due to the birefringence of the molecules of liquid crystal. This method [11]. however. has a disadvantage that the color composition is fully determined by the liquid crystal display structure and the orientation of polarizers.

In this paper. we demonstrate a new method similar to the halftone screen technique [2]. The LCTV generates two distinctive primary color images with contrast reversed. The superposition of the two images produce a pseudocolor encoded

pattern. By changing the primary color filter. we may easily alter the color composition of the output.

2. — PSEUDOCOLOR ENCODING

We now describe a white-light intensity pseudocolor encoding technique for monochrome images. We assume that an intensity distribution is available for pseudocolor encoding. If a LCTV can generate positive and negative images of the object. we may encode the two images with two different primary colors. *Figure 1* shows a sketch of these normalized irradiance as a function of gray scale.

Suppose that the positive and negative image are encoded by red and green filter respectively. When input gray scale increase from 0 to 8. the output color will gradually change from green to yellow to red. To obtain a full-color encoding. this method may be extended to three primary colors as shown in *figure 2.* where the LCTV has to generate an additional intermediate image.
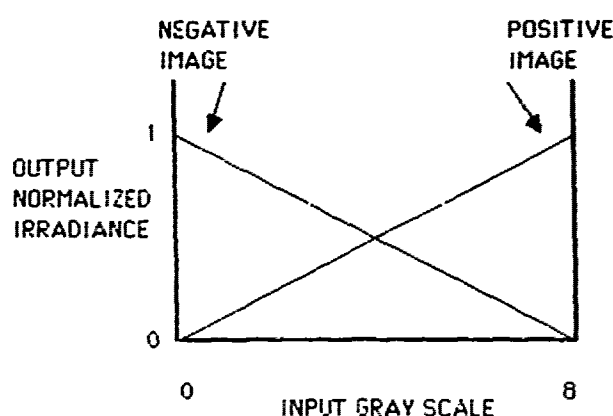
FIG. 1 — *Normalized irradiances as functions of gray scale variation for two-primary-color-encoding.*
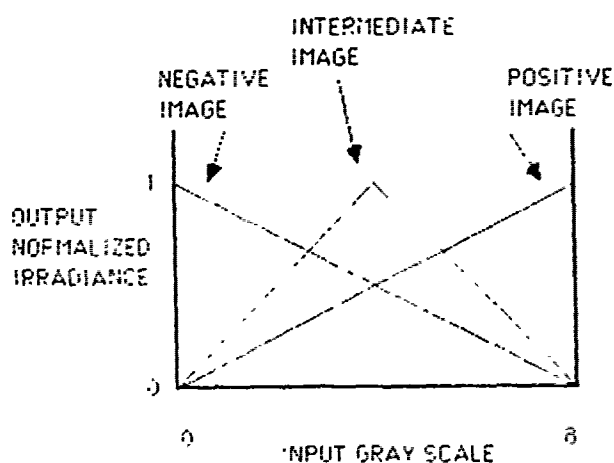
FIG. 2 — *Normalized irradiances as functions of gray scale variation for three-primary-color-encoding*

3. — LIQUID CRYSTAL TELEVISION

A great number of papers have been published based on the work using LCTVs. Only the properties of the LCTV which are important to this experiment will be discussed here. *Figure 3* shows the 90° twisted nematic LC molecules sandwiched between a pair of parallel polarizer and analyser. When there is no voltage applied across the cell. the plane of polariz-
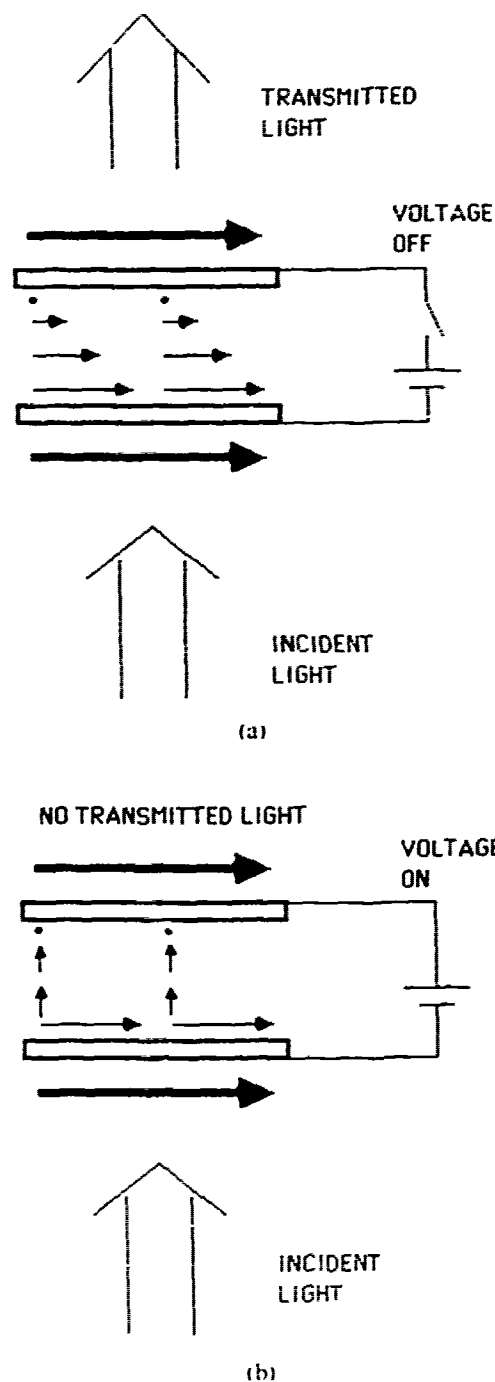

FIG. 3 — *The 90° twisted nematic LC molecules sandwiched between a pair of parallel polarizer and analyser (polarization direction shown in large arrows) The LC molecules alignment (shown in small arrows) changes from (a) when no voltage is applied to (b) when voltage is fully applied across the LC cell*

ation of the incident linear polarized light is rotated through 90° by the LC material, hence no light passes through the analyser. When the voltage across the cell is fully applied, the LC molecules will align in the direction of the applied electric field, hence the polarization direction of light passed through is not rotated and there is total transmission of light through the analyser. If this parallel analyser generates a positive image, the orthogonal analyser would produce a negative image. Chao and Liu [12] reported a pseudocolor technique that superimposed red, green and blue color-encoded images obtained with parallel, 45° and 90° oriented analysers respectively.

The difference between the proposed technique here and that of Chao and Liu [12] is that we utilize the hybrid field effect of the liquid crystal material [13] (twisted nematic field effect and birefringence) while they basically utilized the twisted nematic field effect only. Such a difference can be observed easily from the experimental setup, in which we used a common analyser and variable polarizers, while Chao and Liu used variable analysers and a common polarizer.

The voltage on each pixel is determined by two factors : a brightness control switch controls the bias voltage across all the pixels on the screen and the video input signal controls the voltage across individual pixels. Neglecting the detailed analysis of the physical phenomenon of birefringence, we may develop an empirical model as follows. When input signal is applied to a pixel, the LC molecules shows the birefringence [11]. The polarization axis of incident light is rotated by an angle of $90° - \theta_b - \theta_{on}$ where θ_b is a function of the bias voltage and θ_{on} is a function of the input signal.

The two operating modes of the LCTV we have used in the experiment are as shown in *figure 4*. In the first mode as shown in *figure 4(a)*, the irradiance transmitted is given by

$$I_1 = \cos^2 (90° - \theta_b - \theta_{on}) . \qquad (1)$$

In the second mode as shown in *figure 4(b)* the polarizer $P2$ is rotated by an angle of $- \theta_{on\,max}$, where $\theta_{on\,max}$ is the maximum angle of θ_{on} that the LCTV can achieve. In this experiment (Radio Shack Vision 3 LCTV was used) we found $\theta_{on\,max} \approx 13°$. The transmitted irradiance of the second mode is given by

$$I_2 = \cos^2 (90° - \theta_b - \theta_{on} - \theta_{on\,max}) . \qquad (2)$$

Suppose $\theta_b < \frac{1}{2} \theta_{on\,max}$, $\theta_{on\,max} = 13°$ and $\theta_{on} = 0°$,
then

$$I_1 < I_2 . \qquad (3)$$

(a)

(b)
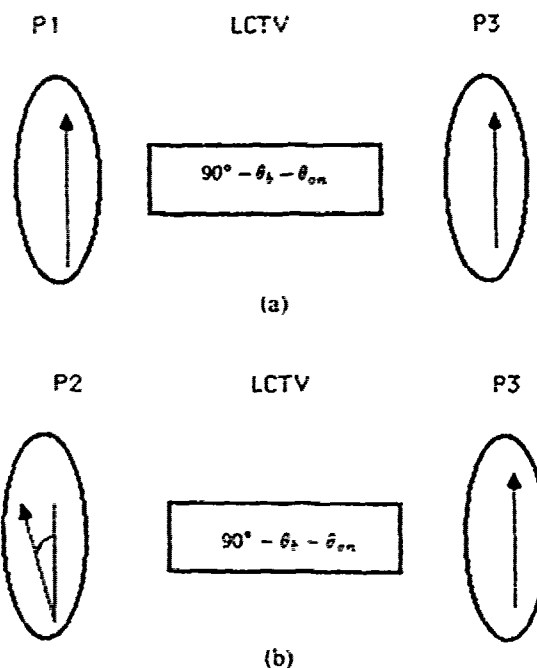
Fig. 4 — *The two operating modes of the LCTV used in the experiment. P1, P2, P3 are polarizers. The arrows show the directions of the polarizers.*

where

$$I_1 = \cos^2 (90° - \theta_b) = \text{dark} . \qquad (4)$$
$$I_2 = \cos^2 (90° - \theta_b - 13°) = \text{bright} . \qquad (5)$$

When $\theta_{on} = \theta_{on\,max} = 13°$.

$$I_1 > I_2 . \qquad (6)$$

where

$$I_1 = \cos^2 (90° - \theta_b - 13°) = \text{bright} . \qquad (7)$$
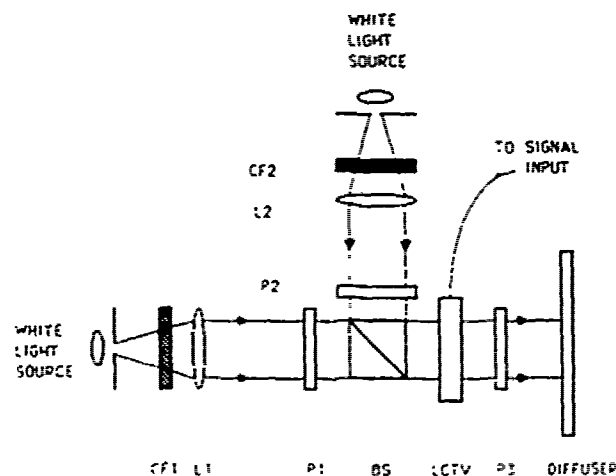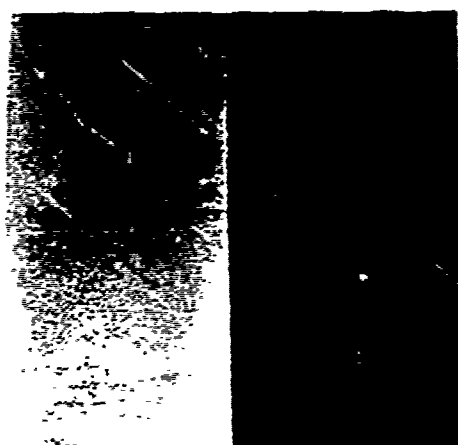$$I_2 = \cos^2 (90° - \theta_b) = \text{dark} . \qquad (8)$$

Fig. 5 — *Optical architecture of gray level pseudocolor encoding. L1, L2 are collimating lenses. CF, CF are color filters. P1, P2, P3 are polarizers. BS is a beam splitter.*

(a)   (b)
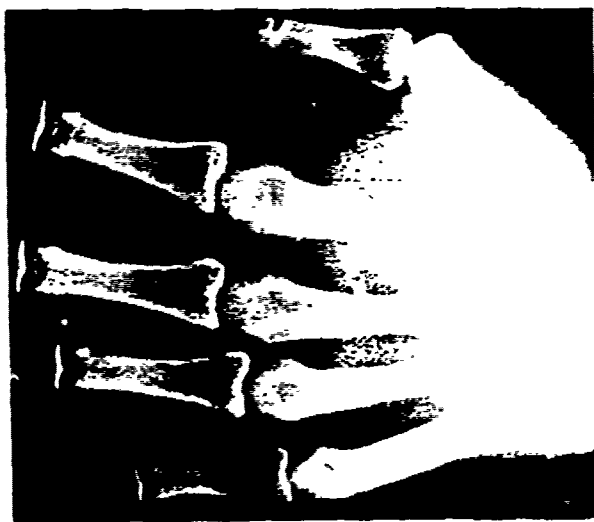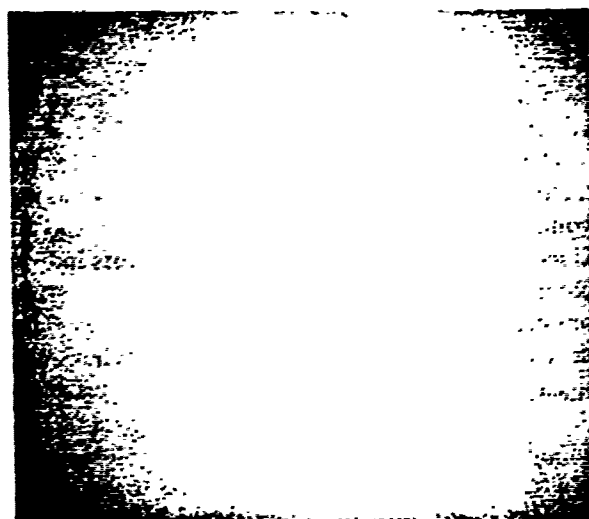
FIG. 6 — (a) *A three-level gray scale chart used as a test object*. (b) *the peudocolor encoded result.*

(a)   (b)

FIG. 7 — (a) *An X-ray transparency used as another test object*. (b) *the pseudocolor encoded result.*

It can be seen that if θ_s is sufficiently small $\left(\theta_s \ll \frac{1}{2} \theta_{s-s} \right)$. then the first and second mode generate high contrast positive and negative image respectively. In other words. the transmitted irradiances given in the two modes are in reversed contrast.

4. — EXPERIMENTS

Indeed. all spatial light modulators based on light polarization modulation. such as magneto-optic spatial light modulator. liquid crystal light valve [13]. micro-channel spatial light modulator. in addition to

the LCTV. can be used in this method. We select to use the LCTV in this experiment mainly due to its features. namely. electrically addressing and low-cost.

The optical setup is as shown in *figure 5*. The combinations of polarizers $(P1. P3)$ and $(P2. P3)$. as discussed in the previous section. simultaneously generate a positive image and a negative image. which are then superimposed on the diffuser. Two distinctive primary color filters CF_1 and CF_2 are positioned at the two optical paths respectively. The pseudocolor-encoded image can then be observed on the diffuser behind the polarizer P3.

A preliminary experiment has been performed to verify the proposed method. In the experiment.

CF_1 was a red filter and CF_2 was a green filter. $P1$ and $P2$ were orientated at $0°$ and $-13°$ with respect to $P3$ accordingly. We first used a three-level gray scale chart as a test object. When the green light was blocked. a positive red image was observed. Similarly. a negative green image was observed by blocking the red light. The gray scale chart and the color encoded output are shown in *figure 6*. Three distinctive colors were observed corresponding to the three gray levels. Another result of a X-ray transparency is demonstrated in *figure 7*. However. the intermediate color (yellow) was dim compared with the red and green color. This might be caused by the low contrast of the original pattern generated on the LCTV.

5. — CONCLUDING REMARKS

A gray level pseudocolor encoding technique that utilizes modulation of polarization in a spatial light modulator has been discussed. Though the proposed technique is applicable to many SLMs such a MOSLM. LCLV. MCSLM. etc.. a preliminary experiment has been demonstrated using a LCTV. The main advantage of this method is that it is an economic real time technique with simple optical architecture. However. with the LCTV we have worked with. the low contrast ratio of the device limits the dynamic range of color content. Also. there is a loss in details of the image due to the quantization of gray level and the limited resolution of the LCTV. With further improvement in the quality of LCTVs. this real time pseudocolor encoding technique may be very useful in medical and industrial applications.

ACKNOWLEDGMENT

REFERENCES

[1] Yu (F. T. S.). — *White Light Optical Signal Processing*. Wiley-Interscience. New York. 1985.

[2] Liu. (H. K.) and Goodman. (J. W.). — " A new coherent optical pseudocolor encoder ". *Nouv. Rev. Opt.*, 7. 285 (1976).

[3] Tai. (A.). Yu. (F. T. S.) and Chen. (H.) — " White-light pseudocolor density encoder ". *Opt. Lett.*, 3. 190 (1978).

[4] Santamaria. (J.). Gea. (M.) and Bescos. (J.) — " Optical pseudocoloring through contrast reversal filtering ". *J. Opt.* 10. 151 (1979).

[5] Chao. (T. H.). Zhuang. (S. L.) and Yu. (F. T. S.) — " White-light pseudocolor density encoding through contrast reversal ". *Opt. Lett.*, 5. 230 (1980).

[6] Mendez. (J. A.) and Bescos. (J.) — " Gray level pseudocoloring with three primary colors by a diffraction grating modulation method ". *J. Opt.* 14. 69 (1983).

[7] Yu. (F. T. S.). Chen. (X. X.) and Chao. (T. H.) — " White-light pseudocolor encoding with three primary colors ". *J. Opt.* 15. 55 (1984).

[8] Yu. (F. T. S..) Chen. (X. X.) and McClure. (K. E.) — " Computer generated linear tricolor sampling pattern and its application ". *J. Opt. (Paris)*. 17. 270 (1986).

[9] Liu. (H. K.) — " Creative applications of liquid crystal television as a spatial light modulator ". CLEO. 1987. paper WJ1.

[10] Yu. (F. T. S.). Jutamulia. (S.) and Hung. (X. L.) — " Experimental application of liquid crystal television to white-light optical processing ". *Appl. Opt.*. 25. 3324 (1986).

[11] Yu. (F. T. S.). Jutamulia. (S.). Lin. (T. W.) and Huang. (X. L.) — " Real-time pseudocolor encoding using a low-cost LCTV ". *Opt. Laser Tech.*. 19. 45 (1987).

[12] Chao. (T. H.) and Liu. (H. K.) — " Optical pseudocolor image enhancement with real-time large screen display ". *Opt. Eng.*. 27. 393 (1988).

[13] Grinberg. (J.). Jacobson. (A. D.). Bleha. (W. P.). Miller. (J.). Fraas. (L.). Boswel. (D.) and Myer. (G.) — " A new real-time non-coherent to coherent light image converter. the hybrid field effect liquid crystal light valve ". *Opt. Eng.*. 14. 217 (1975).

APPENDIX 10.7

White-Light JTC

White-light joint-transform correlator

F. T. S. Yu and Y. S. Cheng

Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

The spatial and temporal coherence requirements in the formation of the joint-transform hologram and in the read-out process are described, and experimental results are given.

The joint-transform correlator (JTC) was proposed by Weaver and Goodman[1] for optically convolving a pair of two-dimensional functions without using a complex spatial filter in the frequency plane. Thus the alignment problem of the spatial filter in the traditional Vander Lugt system[2] is avoided. Recently, by incorporating some spatial light modulators in the optical system, the JTC was applied to real-time pattern recognition.[3] Because there are many advantages, including a higher signal-to-noise ratio, associated with the white-light optical processing system,[4] we investigate here the possibility of utilizing white-light illumination in the JTC.

Consider the JTC shown in Fig. 1, in which Fig. 1(a) is the optical system for recording the joint-transform hologram (JTH) and Fig. 1(b) is the readout optical system. Usually these two systems are connected using a beam combiner. Light emitted by the source point in plane P_0 is collimated by lens L_1 to shine on the objects (both reference and input objects) on plane P_1. A square-law detector is placed in frequency plane P_2 to record the joint Fourier-transform spectrum of the input objects and the reference object. Let a point source at (x_s, y_s) in the source plane illuminate the objects with amplitude transmittance

$$t(x_1, y_1) = a_r(x_1 - x_r, y_1 - y_r) + a_0(x_1 - x_r - X, y_1 - y_r - Y). \quad (1)$$

where a_r is the reference object and a_0 is the input object to be detected. A phase grating G_1 with its line structure parallel to the x_1 axis is placed in contact with the object. A first-order joint-transform spectrum (JTS) is recorded. If we assume linear recording, the amplitude transmittance of the recorded JTH can be written as

$$t_s(x_2, y_2) = |A_r[\cdot]|^2 + |A_0[\cdot]|^2 + 2|A_r[\cdot]||A_0[\cdot]|$$

$$\times \cos\left\{2\pi \frac{X}{\lambda F_2}(x_2 + F_2 x_r/F_1) + 2\pi \frac{Y}{\lambda F_2}\right.$$

$$\left. \times (y_2 + F_2 y_r/F_1 - \lambda F_2 f_1) + \Phi_r[\cdot] + \Phi_0[\cdot]\right\}. $$

$$(2)$$

where $A_r = |A_r|\exp(i\Phi_r)$ and $A_0 = |A_0|\exp(i\Phi_0)$ are the Fourier transforms of a_r and a_0. $[\cdot]$ denotes $[(x_2 + F_2 x_r/F_1)/\lambda F_2, (y_2 + F_2 y_r/F_1 - \lambda F_2 f_1)/\lambda F_2]$. λ is the wave-

length of the light, and f_1 is the spatial frequency of the grating. From Eq. (2) we note that the spectrum is dispersed into a rainbow color. For the wavelength λ, the JTS is centered at $x_2 = -F_2 x_r/F_1$ and $y_2 = -F_2 y_r/F_1 + \lambda F_2 f_1$.

In order to record the JTS successfully, certain coherence requirements should be met. The spatial coherence requirement is discussed first. For simplicity, if we assume that the input object and the reference object are identical, Eq. (2) becomes

$$t_s(x_2, y_2) = 2|A_r[\cdot]|^2\left\{1 + \cos\left[2\pi \frac{X}{\lambda F_2}(x_2 + F_2 x_r/F_1)\right.\right.$$

$$\left.\left. + 2\pi \frac{Y}{\lambda F_2}(y_2 + F_2 y_r/F_1 - \lambda F_2 f_1)\right]\right\}. \quad (3)$$

Let a line, which is parallel to the line going through the center of both objects, in the source plane be called a parallel line. Suppose that the length of projection of the spatially incoherent source on the parallel line is L_r. Summing the contributions from the whole illuminating source, the fringe pattern would be wiped out if the pattern produced by the source point at $L_s/2$ is laterally shifted by one fringe period with respect to that produced by the source point at $-L_s/2$. This
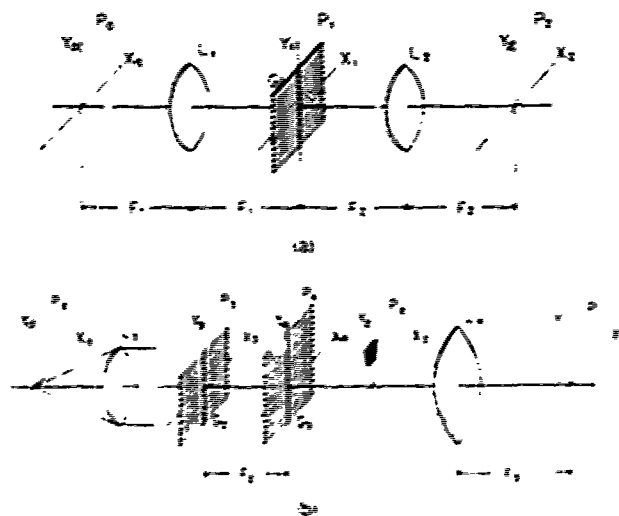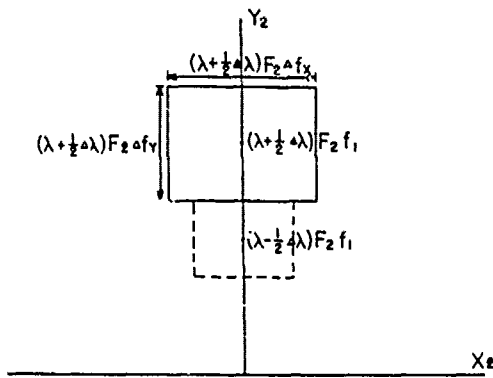

Fig. 1. (a) Optical system for recording the JTH. (b) The readout optical system.

Fig. 2. JTS for estimation of $\Delta\lambda$.

corresponds to the condition $(X^2 + Y^2)^{1/2} = \lambda/(L_s/F_1)$. Since $\lambda/(L_s/F_1)$ is the coherence length in the direction of the parallel line, we note that, in order to produce the fringe pattern with considerable contrast, the separation between two objects should be much smaller than the lateral coherence length.

For simplicity in discussing the temporal coherence requirement, we assume that a white-light point source, at the origin of the source plane, is illuminating the objects. In this condition, Eq. (3) can be simplified t read

$$t_J(x_2, y_2) = 2|A_r[\cdot]|^2 \left\{ 1 + \cos\left[2\pi \frac{X}{\lambda F_2} x_2 \right.\right.$$
$$\left.\left. + 2\pi \frac{Y}{\lambda F_2} (y_2 - \lambda F_2 f_1) \right] \right\}. \quad (4)$$

Suppose that the spatial bandwidth of the object in the x_2 direction is Δf_x and that in the y_2 direction is Δf_y. Referring to Fig. 2, we can estimate the spectral bandwidth $\Delta\lambda$ that contributes to the final fringe pattern. For the spatial spectrum of the object due to $\lambda + \frac{1}{2}\Delta\lambda$ to be separated from that due to $\lambda - \frac{1}{2}\Delta\lambda$, we have

$$\Delta\lambda = \lambda \Delta f_y / f_1, \quad (5)$$

where $\Delta f_y \ll f_1$ is assumed. Since the fringe pattern for each wavelength λ is centered at $x_2 = 0$ and $y_2 = \lambda f_1/F_2$, the dispersion distance $\Delta\lambda F_2 f_1$ should be smaller than the period $(\lambda F_2/Y)$ of the fringe pattern in the y_2 direction for the final fringe pattern not to be wiped out completely. This leads to the criterion

$$Y \Delta f_y < 1, \quad (6)$$

where Eq. (5) is utilized. For example, of an object $\Delta f_y = 10$ mm^{-1}, the off-axis distance Y should be smaller than 0.1 mm. Thus, to have a fringe pattern with high contrast, the centers of both the input and the reference objects should lie on a line parallel to the x_1 axis. In this condition, Eq. (4) becomes

$$t_J(x_2, y_2) = 2|A_r[\cdot]|^2 \left[1 + \cos\left(2\pi \frac{X}{\lambda F_2} x_2 \right) \right], \quad (7)$$

where $[\cdot]$ denotes $[x_2/\lambda F_2, (y_2 - \lambda F_2 f_1)/\lambda F_2]$. To estimate the number of fringes that are available for the readout process, assume that the object has a relatively flat spatial spectrum. Integrating over the spectrum from $\lambda - \frac{1}{2}\Delta\lambda$ to $\lambda + \frac{1}{2}\Delta\lambda$, the cosine term in Eq. (7) becomes

$$\int_{\lambda - \frac{\Delta\lambda}{2}}^{\lambda + \frac{\Delta\lambda}{2}} \cos\left(2\pi \frac{X}{\lambda F_2} x_2 \right) d\lambda$$

$$= \Delta\lambda \, \mathrm{sinc}\left(\frac{\Delta\lambda X}{\lambda^2 F_2} x_2 \right) \cos\left(2\pi \frac{X}{\lambda F_2} x_2 \right), \quad (8)$$

where $\Delta\lambda \ll \lambda$ or $\Delta f_y \ll f_1$ is assumed. From Eq. (8), we find that the number of fringes under both the main lobe and the first sidelobes of the sinc function is $4\lambda/\Delta\lambda$ or $4f_1/\Delta f_y$. For $f_1 = 80$ mm^{-1} and $\Delta f_y = 10$ mm^{-1}, the number of fringes would be 32.

In the readout process the optical system shown in Fig. 1(b) is used. White light from a point source at the origin of plane P_0' is collimated by lens L_3. Two identical gratings, of spatial frequency f_1, are situated on planes P_3 and P_4 with their line structures parallel to the x axis. A rectangular aperture $\mathrm{rect}(x_3/L_x)\mathrm{rect}(y_3/L_y)$ on plane P_3 limits the spatial extent of the plane wave. Owing to the dispersion effect of grating G_2, the first-order wave of each wavelength λ would appear in the region $\mathrm{rect}(x_4/L_x)\mathrm{rect}(y_4 - \lambda F_2 f_1)/L_y)$ on plane P_4, where the diffraction effect on the rim of the rectangle is neglected. When it is incident upon grating G_3, this wave is redirected. The wave, propagating in the direction parallel to the optical axis retained to shine on the JTH. Let the readout w completely overlap the JTS. At plane P, the intensity distribution of the correlation spots would be

$$I_\pm(x, y) = \int_{\lambda_s}^{\lambda_L} d\lambda \left| \iint_{-\infty}^{\infty} \mathrm{rect}\left(\frac{x_2}{L_x} \right) \mathrm{rect}\left(\frac{y_2 - \lambda F_2 f_1}{L_y} \right) \right.$$

$$\times \left[\int_{\lambda\left(1 - \frac{\Delta f_y}{2f_1}\right)}^{\lambda\left(1 + \frac{\Delta f_y}{2f_1}\right)} \left| A_r\left(\frac{x_2}{\lambda' F_2}, \frac{y_2 - \lambda' F_2 f_1}{\lambda' F_2} \right) \right|^2 \right.$$

$$\left. \times \exp\left(\pm i2\pi \frac{X}{\lambda' F_2} x_2 \right) d\lambda' \right] \exp\left(-i2\pi \frac{x}{\lambda F_3} x_2 \right)$$

$$\left. \times \exp\left(-i2\pi \frac{y}{\lambda F_3} y_2 \right) dx_2 dy_2 \right|^2, \quad (9)$$

where I_+ (I_-) denotes the intensity distribution of the
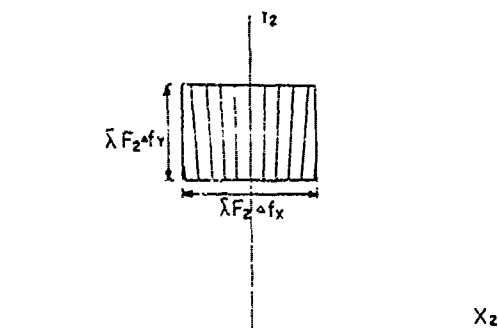
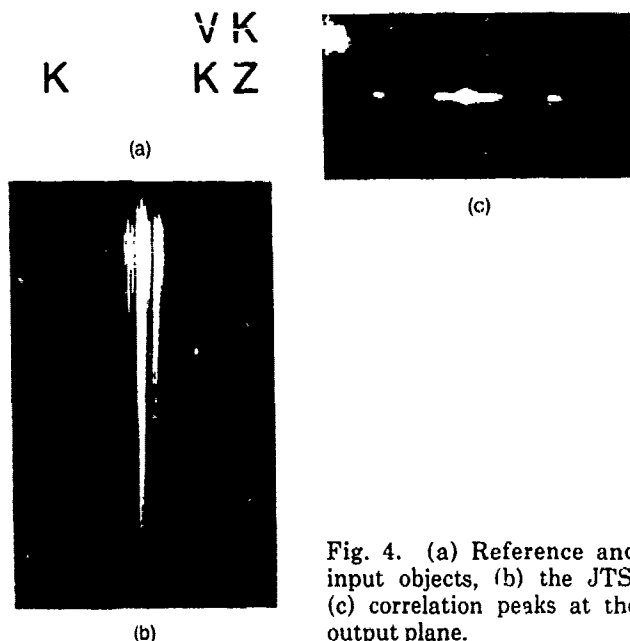Fig. 3. JTH illuminated by a wave with a wavelength λ

Fig. 4. (a) Reference and input objects, (b) the JTS, (c) correlation peaks at the output plane.

desired correlation spot in the region where $x > 0$ ($x < 0$), λ_L is the longest wavelength, λ_S is the shortest wavelength, and Eq. (7) is used. If a source of finite size is utilized instead, the output intensity distribution would be the convolution of Eq. (9) with the source distribution. Equation (9) is complicated and needs further investigation. Here, we give a heuristic explanation. Figure 3 shows the JTH illuminated by a wave with wavelength λ. We take $L_x = \bar{\lambda} F_2 \Delta f_x$ and $L_y = \bar{\lambda} F_2 \Delta f_y$, where $\bar{\lambda}$ is the average wavelength. Owing to the finite size of the readout wave, the linear size of the resolution element is $\lambda F_3/(\bar{\lambda} F_2 \Delta f_x)$ in the x direction. Similarly, it is $\lambda F_3/(\bar{\lambda} F_2 \Delta f_y)$ in the y direction. Since the spatial period of the fringes in the x_2 direction varies from $(\lambda + \frac{1}{2}\bar{\lambda}\Delta f_y/f_1)F_2/X$ for the top of the JTH to $(\lambda - \frac{1}{2}\bar{\lambda}\Delta f_y/f_1)F_2/X$ for the bottom of the JTH in Fig. 3, the center of the desired correlation peak would be at $x = XF_3/F_2$, with a slight broadening $XF_3\bar{\lambda}\Delta f_y/(F_2 f_1)$ owing to nonuniformity of the fringes. Generally, the combined broadening effect due to both nonuniformity of the fringes of the JTH and diffraction is much smaller than the distance between the correlation peaks. For instance, if $\Delta f_x = \Delta f_y = 10$ mm^{-1} and $F_2 = F_3$, the linear size of the correlation spot due to the combined effect is approximately 0.2 mm. Hence, after summing the contributions from all the spectral elements, the size of the white-light correlation spot stays comparable with that due to the wave with average wavelength $\bar{\lambda}$.

In the experiment a mercury-arc source is imaged by a condenser lens onto a pinhole of diameter 15 μm. Since the focal length of the collimating lens is 30 cm. the lateral coherence length would be 11 mm. The letters in Fig. 4(a) are photoreduced on Kodak SO-253 film and become transparent. On the left-hand side, the letter K serves as the reference object, while the four letters on the right-hand side are the input objects. The letters K, K, and Z are aligned to lie on a line parallel to the x_1 axis. The distance between the reference object and the letter Z is 4.5 mm, which is smaller than half the lateral coherence length. In order to meet the requirement $\Delta f_y \ll f_1$, the spatial frequency of the phase grating G_1 is taken to be 80 mm^{-1}. The focal length of lens L_2 is 60 cm. Figure 4(b) shows the JTS in which the fringes under both the main lobe and the first sidelobes of the sinc function are approximately 30 lines. In the readout process, the diameter of the white-light source is 0.2 mm, and the focal length of the collimating lens L_3 and the Fourier-transform lens L_4 is 30 cm. Parallel gratings G_2 and G_3 are identical to grating G_1 and are separated by distance F_2. Shown in Fig. 4(c) is the intensity distribution of the output at plane P. The bright dots on each side of the figure are the correlation peaks of the reference letter K and the input object K [to the left of letter Z in Fig. 4(a)]. There is no output for the letter K that is above the letter Z because relation (6) is severely violated. To identify the letter K that is above the letter Z, we have to rotate the input plane (or grating G_1) about the optical axis until the line going through the center of the reference object and that of the input object K is parallel to the line structure of grating G_1. This suggests the possibility of real-time two-dimensional identification of the input objects. When spatial light modulators are placed in planes P_1 and P_2, a two-dimensional correlation pattern can be produced with bright correlation spots at their corresponding positions if the input plane and the output detector are rotated synchronously.

In summary, we have demonstrated both theoretically and experimentally the possibility of utilizing white-light illumination in the JTC. When the separation between the reference object and the input object is much smaller than the lateral coherence length, and the line joining both objects is parallel to the line structure of the grating, the JTH will have good fringe contrast. By increasing the spatial frequency of the grating, the broadening effect of the correlation spots due to nonuniformity of the fringes can be kept small in the readout process. Since each readout process only gives the information about one line in the object plane, in order to obtain a two-dimensional correlation map, both the input and the output planes should be continuously rotated.

References

1. C. S. Weaver and J. W. Goodman. Appl. Opt 5, 1248 (1966).
2. A. Vander Lugt, IEEE Trans. Inf. Theory IT-10, 139 (1964).
3. F. T. S. Yu and X J. Lu, Opt. Commun. 52, 10 (1984).
4. F. T. S. Yu, White-Light Optical Signal Processing (Wiley, New York, 1985), p. 25.

APPENDIX 10.8

High-Efficient JTC

High-efficiency joint-transform correlator

Francis T. S. Yu

Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

Eddy C. Tam

Department of Physics, University of North Carolina at Charlotte, Charlotte, North Carolina 28223

Don A. Gregory

U.S. Army Missile Command, Research Directorate, Redstone Arsenal, Alabama 35898-5248

A technique that increases the efficiency of a conventional joint-transform correlator is proposed. By modifying the joint-transform power spectrum of the input objects by using a spatial sampling method, the readout light as well as the physical area of the square-law detector can be fully utilized. As a result, the output correlation intensity can be substantially increased. Cases of using coherent and partially coherent readout light are discussed, and experimental results are presented.

In a joint-transform correlator[1,2] (JTC) the spectra of two input image functions $f(x, y)$ and $g(x, y)$ are first recorded by a square-law detector (e.g., a liquid-crystal television, a microchannel spatial light modulator, a charge-coupled-device camera, or photographic film). A plane wave is then applied to read out the joint-transform power spectrum, producing the following complex light field at the output plane:

$$u(x, y) = f(x, y) \otimes f(x, y) + g(x, y) \otimes g(x, y)$$
$$+ f(x, y) \otimes g(x + 2x_0, y) + g(x, y) \otimes f(x - 2x_0, y), \quad (1)$$

where $2x_0$ is the separation between the two input functions on the input plane and $\otimes$ denotes the correlation operation. Note that the last two terms represent the cross-correlation functions that are diffracted around $(\pm 2x_0, 0)$. Owing to the limited dynamic range of the square-law detector, the encoded joint-transform power spectrum is usually confined to a small area. Many of the commercially available square-law detectors used in JTC experiments have a much larger recording area, and therefore the detecting devices are not fully utilized. If the collimated readout light beam is reduced into such a small cross-sectional area, the effective width of the correlation peak would be increased owing to convolution between the correlation function and the focal spot of the readout beam. Therefore, to determine the center of a correlation spot accurately, a readout beam of large cross section is preferable. Moreover, under some circumstances the readout light cannot be reduced into such a small cross section, e.g., in a single spatial-light-modulator JTC architecture in which the collimated light is used for both writing and reading. Readout light power that falls onto the unencoded area of the square-law detector would not be utilized, and the efficiency of the JTC system would be reduced.

In this Letter an image sampling technique is pro-posed to increase the efficiency for a given JTC system. We assume that a one-dimensional sampling grating has been synthesized. The corresponding amplitude transmittance function is given by

$$G(x) = \sum_{n=-\infty}^{\infty} \text{rect}\left(\frac{x - nd}{w}\right) = \sum_{n=-\infty}^{\infty} K_n \exp(jnr_0 x),$$
$$(2)$$

where

$$\text{rect}\left(\frac{x - nd}{w}\right) = \begin{cases} 1 & |x - nd| \leq w/2 \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

$$K_n = \frac{w}{d}\frac{\sin(n\pi w/d)}{n\pi w/d} = \frac{\omega}{d}\text{sinc}(nw/d), \quad (4)$$

w and d are the pulse width and the period of the sampling grating, respectively, and $r_0 = 2\pi/d$. Note that the Fourier expansion of the grating is given as the rightmost expression in Eq. (2).

The grating is then inserted at the input plane of the JTC (as shown in Fig. 1) whose input functions are assumed to be two identical images located at $(x_0, 0)$ and $(-x_0, 0)$. The amplitude transmittance at the input plane is then given as

$$T(x, y) = \sum_{n=-\infty}^{\infty} \{K_n[f(x, -x_0, y)$$
$$+ f(x + x_0, y)]\exp(jnr_0 x)\}. \quad (5)$$

The corresponding joint-transform power spectral distribution recorded at the square-law detector is

$$S(p, q) = |\mathcal{F}\{T(x, y)\}|^2 = \sum_{n=-\infty}^{\infty} 2K_n^2 |F(p - nr_0, q)|^2$$
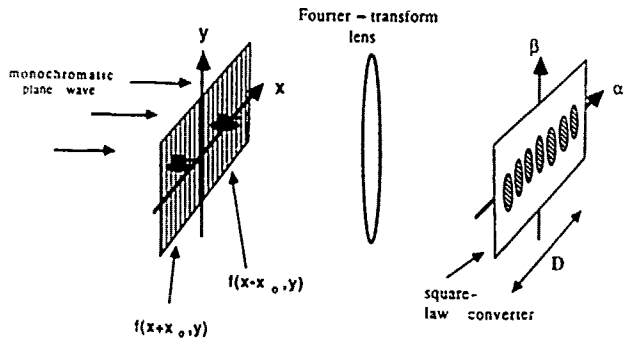$$\times \{1 + \cos[2x_0(p - nr_0)]\}, \quad (6)$$

Fig. 1.   High-efficiency joint-transform architecture.

where $p = (2\pi/f\lambda)\alpha$ and $q = (2\pi/f\lambda)\beta$ are the spatial frequency coordinates, $F(p, q)$ is the Fourier transform of $f(x, y)$, and $\mathcal{F}\{\cdot\}$ denotes a Fourier-transformation process. To ensure that the spectra do not overlap, it is assumed that the sampling rate satisfies the Nyquist requirement, i.e., $d \leq \pi/\Delta p$, where Δp is the angular bandwidth of the input object. Therefore, this determines the number of orders of the spectra that can be written onto the square-law detector. Given this number, to obtain almost equal intensity distribution for all the spectra recorded, the width of the sinc function $\text{sinc}(pw/2\pi)$ should be large compared with the dimension of the detector D, i.e., $D < f\lambda/2w$. Note that the introduction of the modulating term K_n in Eq. (6) produces power spectra of different writing intensities that enhance the dynamic range of the square-law detector.

For a simplified consideration, let us use only the first $\pm N$ orders of diffraction in which the K_n attenuation factor can be ignored, i.e.,

$$S(p, q) \approx \sum_{n=-N}^{N} 2|F(p - nr_0, q)|^2$$

$$\times \{1 + \cos[2x_0(p - nr_0)]\}. \quad (7)$$

If this joint-transform power spectra array is read out with a coherent source, the object irradiance at the output plane becomes

$$I_c(x, y) = \left[1 + 2\sum_{n=1}^{N} \cos(nr_0 x) \right]^2$$

$$\times [|2f(x, y) \otimes f(x, y)|^2 + |f(x, y) \otimes f(x + 2x_0, y)|^2$$

$$+ |f(x, y) \otimes f(x - 2x_0, y)|^2]. \quad (8)$$

By comparing the above results with those obtained from a conventional JTC [i.e., Eq. (1)], it can be seen that the average energy of the correlation function in Eq. (8) is increased by

$$\frac{1}{d} \int_{x=0}^{x=d} \left[1 + 2\sum_{n=1}^{N} \cos\left(\frac{n2\pi}{d} x\right) \right]^2 dx = 1 + 2N \text{ times.}$$

$$(9)$$

However, despite the increase in the average peak intensity, the correlation function is modulated by the square of a sinusoidal factor, and this might not be a desirable effect in some cases.

On the other hand. if quasi-monochromatic, partially coherent light is utilized as the readout source, whose coherence width equals the width of the joint-transform power spectrum, the object irradiance at the output plane can be written as

$$I_{pc}(x, y) = (1 + 2N)[|2f(x, y) \otimes f(x, y)|^2 + |f(x, y)$$

$$\otimes f(x + 2x_0, y)|^2 + |f(x, y) \otimes f(x - 2x_0, y)|^2]. \quad (10)$$

The output correlation function is simply amplified by $1 + 2N$ times that of a conventional JTC. There is no modulation within the correlation function, and maximum intensity at the center of correlation is guaranteed. Moreover, use of partially coherent readout light would also suppress coherent artifact noise and thus increases the signal-to-noise ratio. Note that the above analysis for a one-dimensional sampling grating can be easily expanded to a two-dimensional sampling case. The spectra array given in Eq. (4) will then become a two-dimensional square array in which the multiple orders of the joint-transform power spectrum will be located at (nr_0, mr_0). Following a similar mathematical analysis, which leads to Eq. (10), the correlation spot intensity will be increased by a factor of $(1 + 2N)^2$ times when a two-dimensional sampling grating is used.

To verify our discussion, an experiment using a $2f$ Fourier-transform system as shown in Fig. 1 was performed. For easy implementation, two pieces of commercially available Ronchi rulings were used to produce a two-dimensional sampling grid. The Fourier-transform spectrum of the grating itself shows that the ± 1 orders are approximately half the intensity of the
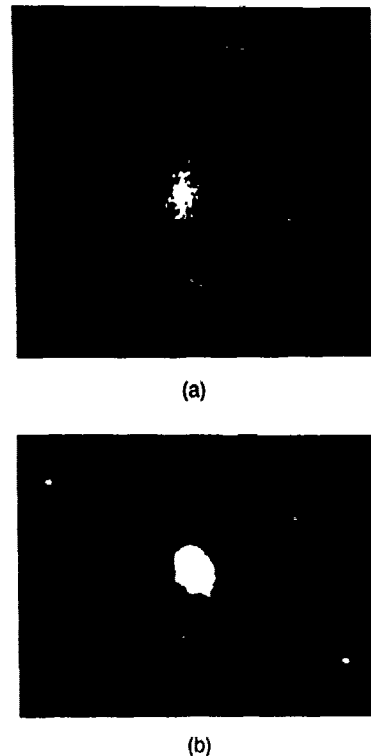

(a)

(b)

Fig. 2.   Experimental results of the JTC when no sampling grating was used on the input plane: (a) joint-transform power spectrum, (b) correlation peak obtained by using a partially coherent readout light.
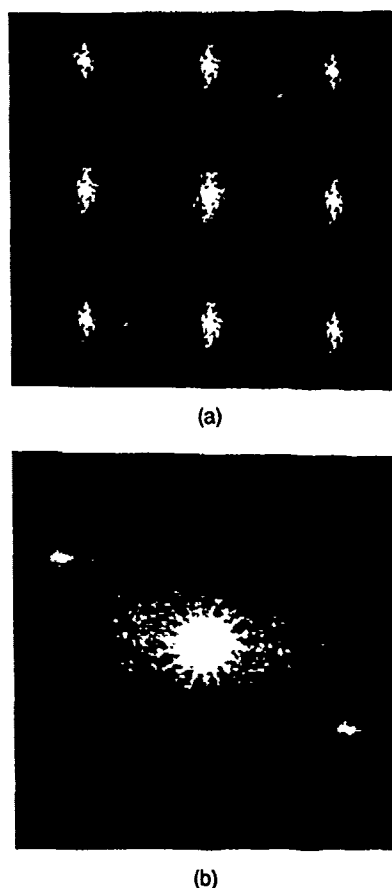
(a)

(b)

Fig. 3. Experimental results of the JTC when a two-dimensional sampling grating was inserted: (a) the 3 × 3 power spectrum array, (b) correlation peaks.

zero order. The higher orders are weak and to permit an easy analysis were not recorded. The input pattern consists of two identical images recorded on 35-mm photographic film. A He–Ne laser was used as the coherent light source, and the power spectrum was recorded on film. Figures 2(a) and 3(a) show the power spectrum recorded before and after the sampling grating was inserted, respectively. Partially coherent light generated from a white-light source in conjunction with an interference filter was then used to read out the power spectrum. Figure 2(b) shows the correlation result when a single spectrum order was used, and Fig. 3(a) shows the result obtained when a 3 × 3 spectra array was used. It was measured that the correlation peak intensity is increased by approximately 4.8 times when the sampling grating is inserted at the input. However, according to the mathematical predication, the peak energy should be increased by 9 times. We believe that the major reason for the decrease in experimental values is because of the weakened modulation in the ±1 orders of diffraction from the Ronchi rulings.

In conclusion, we have proposed and demonstrated an image sampling technique that improves the efficiency for a given JTC system. By writing multiple orders of the joint-transform spectrum onto the square-law detector, the physical size of the device as well as the readout light power are more fully utilized. Consequently, the correlation peak intensity is substantially increased. Under either coherent or partially coherent readout light, the energy of the correlation function thus produced is proportional to the number of orders of spectra used. Such merits of this proposed technique would be helpful under low light intensity. Moreover, it must be mentioned that the increase in the correlation peak intensity with this technique is different from simply increasing the readout light power while reading a single order of the joint-transform spectrum. When random noise is added in during the JTC process, the latter approach would increase any noise spikes present on the output plane by the same amount as for the correlation peak. However, in our proposed technique, each spectrum order being used would produce different noise spikes at different positions. Therefore this technique would also increase the discrimination ability of the JTC under noisy conditions.

Note that if a charge-coupled-device camera connected to a computer is used as the square-law detector, then the sampling gratings need not be used in the optical setup: A simple digital image-processing algorithm can be applied to replicate a single order of the spectrum into multiple orders of the spectral array. Since the JTC's can be regarded as generalized optical processors,[3] the proposed technique can improve the performance in other image-processing applications.

References

1. C. S. Weaver and J. W. Goodman, Appl. Opt. 5, 1248 (1966).
2. F. T. S. Yu and X. J. Lu, Opt. Commun. 52, 47 (1984).
3. F. T. S. Yu and J. E. Ludman, Microwave Opt. Technol. Lett. 1, 374 (1988).

APPENDIX 10.9

Autonomous Target Tracking

Autonomous real-time object tracking with an adaptive joint transform correlator

Eddy C. Tam, MEMBER SPIE
Francis T. S. Yu, FELLOW SPIE
The Pennsylvania State University
Department of Electrical Engineering
University Park, Pennsylvania 16802

Don A. Gregory, MEMBER SPIE
U.S. Army Missile Command
Research Directorate
Redstone Arsenal, Alabama 35898-5248

Richard D. Juday, MEMBER SPIE
NASA
Johnson Space Center
Houston, Texas 77058

Abstract. A hybrid optical/digital system for tracking an object in a sequence of images is described. Since a joint transform correlator does not require a matched spatial filter in the correlation process, object tracking can be carried out by continuously updating the reference image with the object image in the previous frame. This adaptive property of a joint transform correlator, together with the parallelism and high processing speed of an optical system, ensure high correlation between objects in two sequential frames. The relative position of the object can then be determined based on the location of the correlation peak. System performance is evaluated and experimental demonstrations are presented.

Subject terms: *pattern recognition; liquid crystal television spatial light modulator; target tracking; joint transform correlator; optical information processing; hybrid optical correlator.*

CONTENTS

1. INTRODUCTION

Optical pattern recognition has attracted much attention in various applications because of its parallel processing and high-speed operation capabilities, particularly in real-time applications such as missile guidance, vehicle tracking, and automated lander guidance in aerospace missions.[1] In 1974, Vander Lugt[2] demonstrated the use of matched spatial filters (MSFs) in performing cloud motion analysis in a sequence of photographs taken from a satellite in half-hour intervals. A new MSF was made for every photograph in order to correlate with the scene

in the next picture, and therefore this is not a practical real-time technique. With the recent development of various kinds of spatial light modulators (SLMs),[3-7] many real-time tracking techniques have been demonstrated. Gara[8] used a liquid crystal light valve to perform real-time tracking based on a fixed MSF in a Vander Lugt correlator. Using a liquid crystal television spatial light modulator, a dichromated gelatin multifocus holo-lens and a MSF array, Chao and Liu[9] demonstrated a technique for simultaneously tracking multiple objects. Optical tracking using novelty filters and image subtraction have also been proposed.[10,11] in which the system detects the difference between sequential image frames and then computes the position of the object in motion.

The joint transform correlator (JTC)[12,13] is an alternative approach to optical pattern recognition. It has been demonstrated that a real-time programmable JTC can be implemented using two SLMs[14] or just a single liquid crystal television (LCTV) in a video feedback fashion.[15,16] In this paper, a technique for real-time adaptive object tracking based on a JTC is proposed and demonstrated. The basic idea is to correlate the object in the current frame with the object pattern in the previous frame of sequential video images so as to obtain the relative locations of the object in the two frames. It must be pointed out that in a
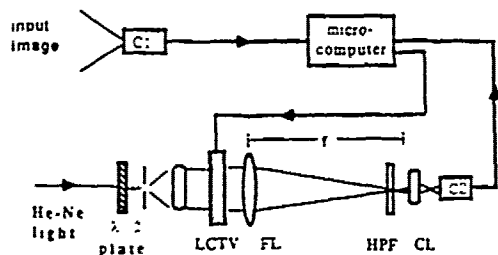
Fig. 1. Optical architecture for the adaptive JTC. C1, C2 are two CCD cameras. FL is a 48 in. Fourier transform lens, and CL is a 50 mm camera lens. The high-pass filter (HPF) is a transparent film with a tiny blurred dot at the center.

Fig. 2. FT spectrum of the LCTV.

Fig. 3. Pixel structure of the color LCTV.

JTC architecture, the reference image and the input scene to be correlated are placed side by side on a spatial light modulator. This unique property of a joint transform correlator enables the system to adapt to the input scene by continuously updating the reference image with the previous frame to correlate with the dynamic object. The key element to this optical JTC is a high-contrast LCTV, and the hybrid system is fully controlled by a microcomputer.

2. OPTICAL ARCHITECTURE AND THEORY

To improve the simplicity and compactness of the system, a single LCTV (Seiko color LCTV model LVD202, 220×240 pixels) together with a Sony CCD camera are used in a video feedback architecture, shown in Fig. 1. Since this is a color LCTV, three neighboring RGB pixels are addressed as one pixel unit when displaying a gray level image, thus decreasing the resolution of the device. By observing its Fourier transform spectrum (Fig. 2), we deduced that the RGB pixels are arranged in the fashion shown in Fig. 3. This is also verified by observing the pixel structure under a microscope. The effect of this structure is discussed in Sec. 4.

To interface this commercial product with our system, the diffuser and the protective glass of the LCTV were first removed. To compensate for the phase distortion, the liquid crystal display (LCD) unit was then separated from the hardware circuit box and immersed in a liquid gate filled with mineral oil. The two original polarizers attached to the display unit were left untouched. A half-wave plate was inserted in front of the He-Ne laser to align the plane of polarization of the linear polarized light with the LCTV's front polarizer. In this arrangement, maximum transmission and a higher contrast ratio can be obtained.

We now discuss the principle of the tracking algorithm. Let us assume that two sequential scenes of a moving object are displayed on the LCTV via the computer's video frame grabber. The previous frame and the current frame are positioned in the upper and lower half of the LCD, respectively, as depicted in Fig. 4. Let

$$f_{t-1}\left(x - x_{t-1}, y - y_{t-1} - \frac{a}{2}\right) \text{ and}$$

$$f_t\left(x - x_{t-1} - \delta x, y - y_{t-1} - \delta y + \frac{a}{2}\right)$$

be the image functions of the objects displayed on the LCD, where $2a$ is the height of the display unit, t and $t - 1$ are the
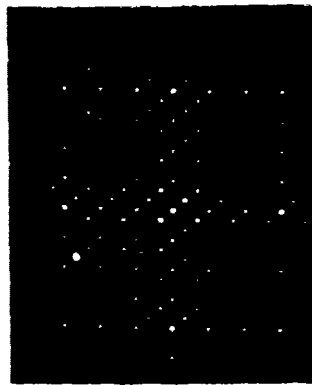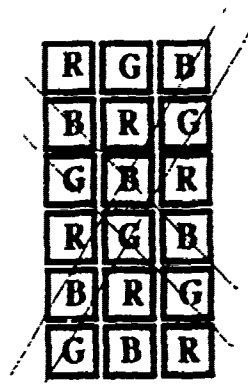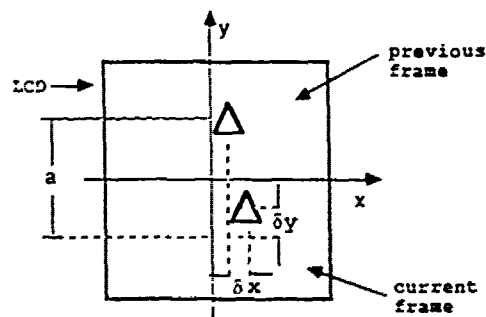

Fig. 4. Arrangement of the two frames on the LCD. $2a$ is the height of the display unit.

current and previous frames, and $(\delta x, \delta y)$ is the relative translation of the object from the $(t - 1)$th to the tth frame. Then the complex light field at the frequency plane of lens FL is given by

$$T(u,v) = F_{t-1}(u,v)\exp\left\{-i2\pi\left[ux_{t-1} + v\left(y_{t-1} + \frac{a}{2}\right)\right]\right\}$$

$$+ F_t(u,v)\exp\left\{-i2\pi\left[u(x_{t-1} + \delta x) + v\left(y_{t-1} + \delta y - \frac{a}{2}\right)\right]\right\}. \tag{1}$$

The power spectrum recorded on CCD camera C1 is then sent back to the LCTV for an inverse Fourier transformation. Thus, the complex light field at the output plane becomes

$$C(x,y) = \mathcal{F}^{-1}\{|T(u,v)|^2\}$$

$$= R_{t,t}(x,y) + R_{t-1,t-1}(x,y)$$

$$+ R_{t,t-1}(x + \delta x, y - \delta y - a)$$

$$+ R_{t-1,t}(x - \delta x, y - \delta y + a). \tag{2}$$

where

$$R_{m,n}(x,y) = \mathcal{F}^{-1}\{F_m(u,v)F_n^*(u,v)\}$$

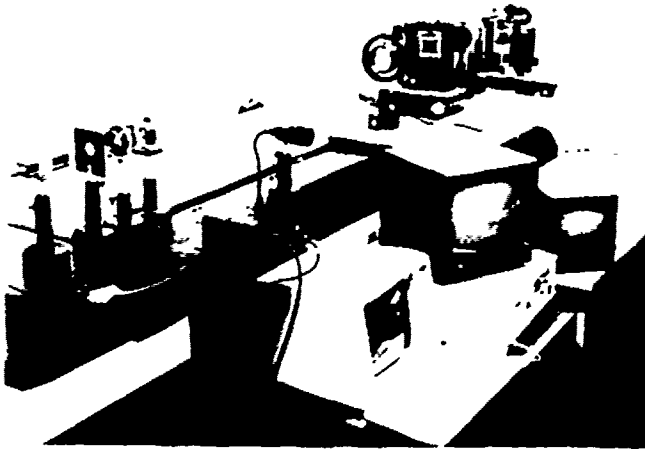$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_m(u,v)f_n^*(u - x, v - y)dudv$$
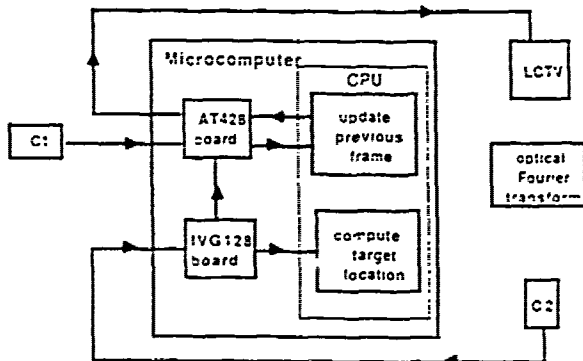
Fig. 5. Experimental setup of the hybrid system.

Fig. 6. Block diagram of the optical digital interface.
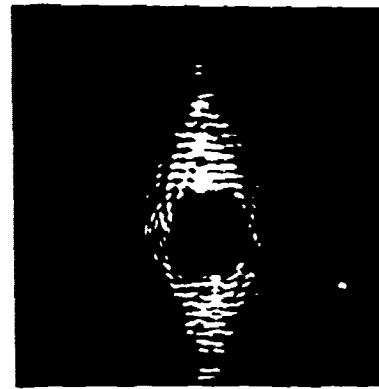

Fig. 7. Joint transform spectrum recorded by C2. The dc light is blocked by the high-pass filter.
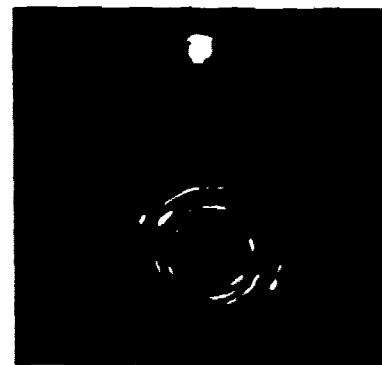
Fig. 8. Correlation result from Fig. 7. One of the two peaks is shown at the top.

is the correlation function between f_m and f_n and $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse Fourier transform of the function $\{\cdot\}$. Note that the first two terms of Eq. (2) are the zero-order terms diffracted around the origin of the output plane, while the last two terms represent the cross-correlation terms diffracted at $x_{peak1} = \delta x$, $y_{peak1} = (\delta y - a)$ and $x_{peak2} = -\delta x$, $y_{peak2} = (-\delta y - a)$. With reference to our previous investigations, the angular and scale tolerance of a target in a JTC is roughly about $\pm 5°$ and $\pm 10\%$, respectively. By assuming that the motion of the object is relatively slow compared to the processing cycle of the correlator, f_{t-1} will strongly correlate with f_t, and two high-intensity correlation peaks can be observed at the output plane. The new location of the object at time t is then given by

$$x_t = (x_{t-1} + x_{peak}), \quad y_t = (y_{t-1} + y_{peak} - a) \tag{3}$$

In our experiment, a 50 mm camera lens was used to magnify the Fourier spectrum to about 5 to 10 lines/mm. A small circular opaque dot with a blurred edge, which acts as a high pass filter, was inserted at the Fourier plane to block the dc light from the CCD, which would otherwise cause multiple reflections and scattering at the surface of the CCD. A photograph of the experimental setup is shown in Fig. 5.

It must be mentioned that this adaptive correlation technique requires the location of the object to be initialized in the first frame. This can be done using the same joint transform correlation system but with a prestored reference image located at a fixed position on the LCTV. In actual applications, the object might have distortion variations due to variations in size, rotation orientation, and perspective. Therefore, a hierarchy search of the object might be needed. To speed up the searching process, distortion invariant filters such as the circular harmonic expansion filters or the synthetic discriminant function filters can be employed.

3. OPTICAL DIGITAL INTERFACE

The hybrid system is controlled by using a C language program, which runs on an IBM AT compatible microcomputer at 10 MHz cpu clock speed. Figure 6 is a block diagram of the digital system hardware configuration. Two video frame grabbers (IVG128 and AT428 from Datacube Inc.), which can digitize a full frame of input video signal into a 384 × 512 image array, have been installed in the microcomputer. The video input of the AT428 board is connected to a Fairchild CCD camera (C1), which is aimed at the object of interest. The video output of the AT428 board is fed to the LCTV via a video rf signal converter. CCD camera C2 is connected to the IVG128 board and captures the joint transform power spectrum during the first half of the processing cycle. Since we do not use a video multiplexer, the recorded spectrum is sent from the IVG board to the AT board
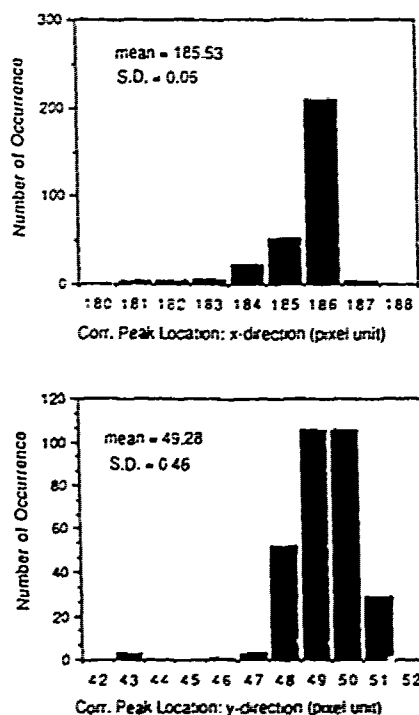
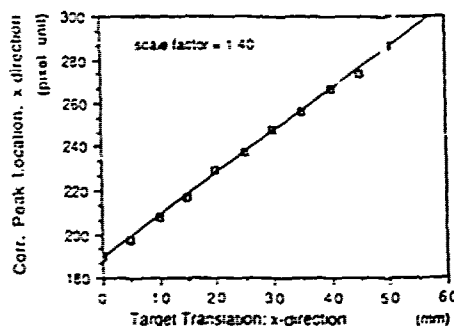Fig. 9. Distribution of the correlation peaks in 300 JTC cycles.

Fig. 10. Correlation peak location as a function of the object's translation in the x direction. The object's overall translation displayed on a 185 mm × 240 mm monitor screen was 68.5 mm. The scale factor was computed in terms of the distance in the monitor space per pixel unit.
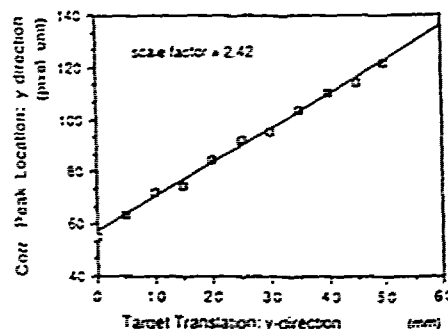
Fig. 11. Correlation peak location as a function of the object's translation in the y direction. The overall translation on the monitor screen was 27.3 mm.

before it can be displayed on the LCTV. During the second half-cycle, the correlation result is captured by C2, and the peak location is determined by sequentially scanning the image array for the pixel with maximum intensity. A position sensitive detector or a parallel array detector would definitely increase the detection speed. Despite these limitations due to the use of general-purpose hardware, our system runs at approximately 1.2 s per cycle. It is possible for the hybrid system to run at half the video frame frequency, which is $\frac{1}{2} \times 30 = 15$ cycles per second, if specialized supporting hardware are used.

Moreover, if an optically addressed spatial light modulator is used as the square law converter instead of a CCD camera and a LCTV used in video feedback fashion, then the processing speed could run at video frame rate, assuming that the response time of the second spatial light modulator is much shorter than a video frame cycle. However, the power consumption and size of a double spatial light modulator architecture would be two times higher and larger than that of this proposed single spatial light modulator architecture.

4. SYSTEM PERFORMANCE

To see the performance of the system, CCD camera C1 was first focused on a stationary object. The JTC then correlated two identical scenes. The joint transform spectrum and correlation output captured by C2 are shown in Figs. 7 and 8, respectively. We slightly modified the program so that the original image displayed is locked into the upper half of the LCD and the updating input frame is always displayed on the lower half. The program was then set to run for 300 cycles, and the distribution of the correlation peak locations is plotted in Fig. 9. The jittery video signal amplified by the low-resolution color pixel structure of the LCTV might be one of the main sources of error that cause fluctuation in the data.

Next, we translated the object along the x direction from 0 to 50 mm in 5 mm increments, and 300 sample data of the correlation peak locations were recorded at every step. Figure 10 shows a plot of the mean of each of the data sets as a function of the object's translation. The result shows an excellent linear relationship between the object space and the correlation space in the x direction. From the plot, the scale factor between the two coordinates in the x direction can be evaluated in terms of the distance in the object space (or in the video system's space) per pixel unit.

The same test was carried out in the y direction, and the result is plotted in Fig. 11. An overall linear relationship is observed, and the y direction scale factor is also evaluated. It is interesting to note that the ratio of these two scale factors,

$$\frac{x \text{ direction scale factor}}{y \text{ direction scale factor}} = \frac{1.401}{2.422} = 0.568 .$$

should be the same as the inverse of the aspect ratio of the image array, given by[18]

$$\frac{1}{\text{aspect ratio}} = \frac{\text{monitor screen height} \times \text{No. horizontal pixels}}{\text{monitor screen width} \times \text{No. vertical pixels}}$$

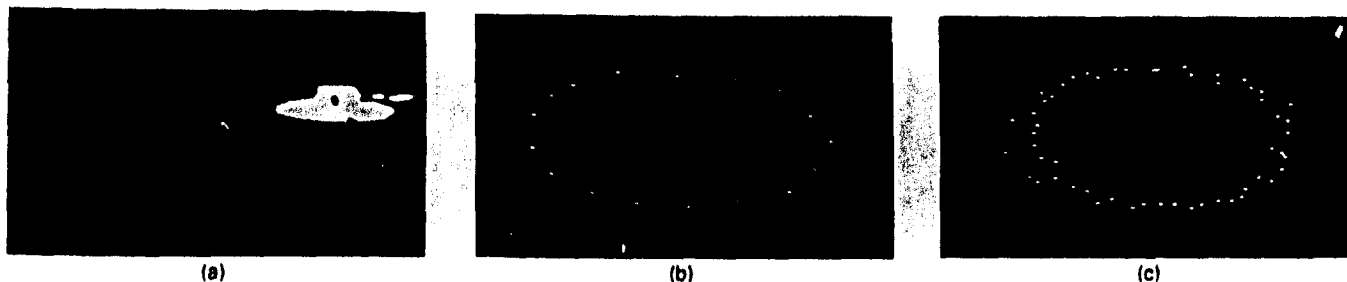which is a measurement of the pixel element's height to width

(a)

(b)

(c)

Fig. 12. (a) A graphical image revolves in an elliptical path at 21 s/rev. The dark dots are the hand track of the tank's positions at each of the JTC cycles. (b) Tracked positions of the tank after one revolution and (c) after four revolutions.
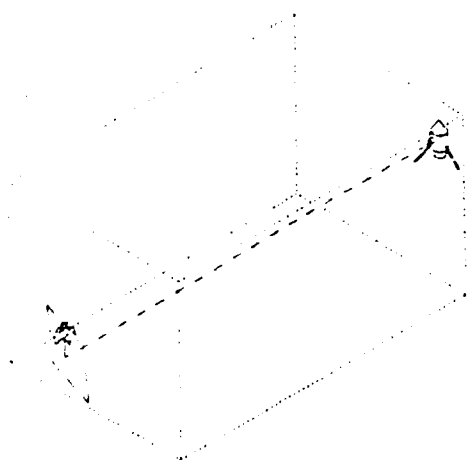

Fig. 13. A camera on a space vehicle aimed at a fixed target on the ground for navigational purposes.

ratio and is equal to 0.578 in our video system (a Panasonic B/W monitor in its under-scan-size mode). Comparing this number with the x,y scale factor ratio, the two values are different by less than 2%. This evaluation can serve as a check of the system's overall accuracy.

A close examination of Fig. 11 shows a periodic stair-like data structure that is absent in Fig. 10, and we have also recorded that the standard deviations at the "steps" are higher than usual. This can be explained as follows: When the object is translated in the x direction, the joint transform fringes are generally aligned at an angle. When the object is translated in the y direction, the fringes are aligned in the x direction. A study of the pixel structure in Fig. 2 reveals that this LCTV is better able to display slanted lines and vertical lines than horizontal lines. The coarse sampling in the horizontal direction thus generates higher data fluctuations and periodic setbacks.

After these preliminary tests, a tank-shaped graphical image displayed on another video monitor was used as the next test object. The tank was set to revolve in an elliptical path at a period of 21 s/rev. [Fig. 12(a)]. The original program, which updates both halves of the LCD, was set to run autonomously. Figures 12(b) and 12(c) show the locations of the tank tracked by the system; proper x and y direction scale factors are used in these plots. Excellent tracking of the object's location is obtained at the end of the first revolution. However, after four revolutions, some deviations of the tracked points from the cor-

rect locus are observed. This reveals that detection error accumulates during each correlation cycle, as can also be seen easily from Eq. (3). This drawback of an adaptive correlation system must be considered and solved by using some other method.
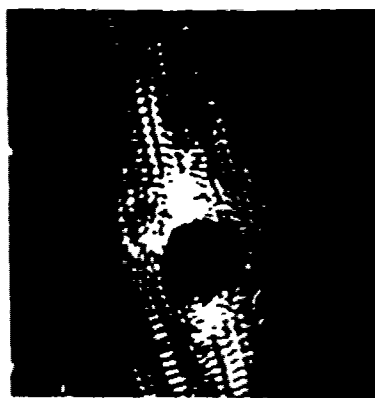
5. ADAPTIVE CORRELATION

A major merit of the proposed technique is the adaptive property of the system. We have carefully simulated two sets of image sequences to illustrate this concept. The first set simulates a situation in which a camera mounted on a moving vehicle, say, a space vehicle, is imaged onto a fixed target on the ground for navigational purposes, as shown in Fig. 13. As the space vehicle approaches the target, the detected scenes are continuously changing; the target size appears larger, and its orientation and shape also change due to different perspective angles from the vehicle in motion. Using computer-aided design graphics, a 3-D tree-like object was created as a target on the ground. Nine image sequences simulating exactly the same scenes as captured by the moving space vehicle were recorded from a CAD terminal, as shown in Fig. 14. The hybrid tracking system has little difficulty correlating the targets between frames, even if the target on the first frame and the last frame look very different. Figure 15 shows the tracked locations of the target as seen from the vehicle's coordinate frame.

Another simulation illustrates a stationary camera/moving object situation. Figure 16 exhibits the motion sequence of a toy tank moving from lower left to upper right in the scene. An overturned truck, which serves as stationary clutter, was also added to the center of the image scene. In this case, four objects are displayed on the LCD during the correlation process, and eight cross-correlation functions are generated on the output plane, four of which are the weak cross-correlations between the tank and the clutter and can be ignored. Based on prior knowledge that the clutter is stationary, the unwanted peaks are always located at $(0, \pm a)$, which can then be distinguished from the peaks generated by the target in motion. Figure 17 shows the joint transform spectrum and the correlation peaks based on frames 1 and 2 of the image sequences. The tracked positions of the target are given in Fig. 18.

6. CONCLUSIONS

In this paper we have demonstrated the use of an adaptive JTC for object tracking based on the correlation between images in sequential scenes. The high-speed parallel processing ability of

Fig. 14. A sequence of nine images were recorded from a CAD terminal, simulating the exact scenes as captured by a camera mounted on the moving space vehicle. Only frames 1, 5, and 9 are shown here.

Fig. 15 Tracked positions of the ground target as seen from the vehicle's coordinate frame

Fig. 16. Motion sequence of a toy tank and a stationary truck. Six frames were taken from a fixed camera; only frames 2, 4, and 6 are shown here. Note that in frame 4, the tank is partially occluded by the clutter

the optical flow is the main advantage of this real time system. Also, the adaptive property of this technique provides data a portion free pattern recognition, which is vital to robust track ing application. The simplicity and compactness of the system properly suggest that it can be implemented in various civilian military, and aerospace applications. We have performed a preliminary analysis and obtained safe factors within a mini mum period of ... per cycle. Nevertheless, this does not reflect the true potential in terms of speed accuracy, and it juestifies that one can the expect from a more sophisticated system implementation this architecture

8. REFERENCES

[references illegible]

(a)

corr. peak from target | stationary corr. peak

(b)

Fig. 17. (a) Joint transform spectrum based on frames 1 and 2. (b) Two correlation peaks appear on the output plane.
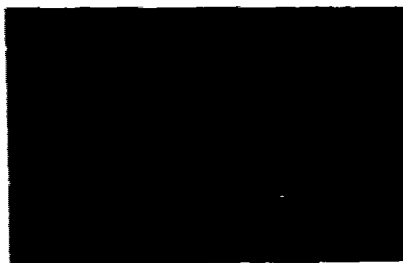

Fig. 18. Tracked positions of the tank in six frames.

15. F. T. S. Yu, S. Jutamulia, T. W. Lin, and D. A. Gregory, "Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator," Appl. Opt. 26, 1370 (1987).
16. B. Javidi, D. A. Gregory, and J. L. Horner, "Single modulator joint transform correlator architectures," Appl. Opt. 28, 411 (1989).
17. F. T. S. Yu, X. Li, E. Tam, S. Jutamulia, and D. A. Gregory, "Detection of rotational and scale varying objects with a programmable joint transform correlator," in Optical Pattern Recognition, H. K. Liu, ed., Proc. SPIE 1053, 167–176 (1989).
18. See for example, Microsoft QuickC Computer Programmer's Guide, p. 101, Microsoft Corp., Redmond, Wash. (1987).

Eddy C. Tam received his BS degree in physics in 1984 from the State University of New York College at Buffalo and his MS degree in physics in 1987 from the Pennsylvania State University, where he is currently a Ph.D. candidate in electrical engineering. His research interests include optical image processing and pattern recognition, holography, and spatial light modulators.

Francis T. S. Yu received his BSEE from Mapua Institute of Technology, Manila, Philippines, and his MS and Ph.D. degrees in electrical engineering from the University of Michigan. From 1958 to 1965 he was a teaching fellow, instructor, and lecturer in the Electrical Engineering Department and a research associate with the Communication Sciences Laboratory at the University of Michigan. From 1965 to 1980 he was on the faculty of the Electrical and Computer Engineering Department at Wayne State University. He was a visiting professor in electrical and computer engineering at the University of Michigan during 1978–1979. Since 1980, Dr. Yu has been a professor in the Electrical Engineering Department at the Pennsylvania State University. He has been a consultant to several industrial and government laboratories. He is active in research in the fields of optical signal processing, holography, optics, and information theory and has published numerous technical papers in these areas. He was a recipient of the 1983 Faculty Scholar Medal for Outstanding Achievement in Physical Sciences and Engineering and the 1984 Outstanding Researcher in the College of Engineering at Penn State and was named Evan Pugh Professor of Electrical Engineering there in 1985. He is the author of Introduction to Diffraction, Information Processing and Holography; Optics and Information Theory; Optical Information Processing; and White-Light Optical Processing. Dr. Yu is a Fellow of OSA, SPIE, and IEEE.

Don A. Gregory was born in Alabama in 1953. He earned BS and MS degrees in physics and a Ph.D. in physics in 1984 from the University of Alabama in Huntsville. Dr. Gregory is currently a research physicist at the U.S. Army Missile Command at Redstone Arsenal, Ala. His interests include optical computing, pattern recognition, and teaching. He is an adjunct professor of physics at the University of Alabama in Huntsville and the University of Alabama in Birmingham. He is also an adjunct professor of electrical engineering at the Pennsylvania State University.

Richard D. Juday holds a BA in physics and a Ph.D. in electrical engineering. Among other areas he has worked in radiation environments for manned space flight and earth resources remote sensing. He currently works in digital and optical image processing for machine vision. Video-rate image coordinate transformations, one aspect of his machine vision activities, are applied to human low vision problems. He has avid interests in whitewater boating, bad minton, and Mexican mountain-stomping and canyon-crawling.

8. A. D. Gara, "Real-time tracking of moving objects by optical correlation," Appl. Opt. 18, 172 (1979).
9. T. H. Chao and H. K. Liu, "Real-time holographic tracking of moving objects," Appl. Opt. 28, 226 (1989).
10. D. Z. Anderson, D. M. Lininger, and J. Feinberg, "Optical tracking novelty filter," Opt. Lett. 12, 123 (1987).
11. Y. Li, A. Kostrzewski, D. H. Kim, and G. Eichmann, "Compact real-time white light optical tracking novelty filter," Opt. 28, (1989).
12. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. 5, 1248 (1966).
13. J. E. Rau, "Detection of differences in real distributions," J. Opt. Soc. Am. 56, 1490 (1966).
14. F. T. S. Yu and X. J. Lu, "A real time programmable joint transform correlator," Opt. Commun. 52, 10 (1984).

APPENDIX 10.10

Image Deconvolution

DEBLURRING USING JOINT TRANSFORM PROCESSOR

F.T.S. YU and Y.S. CHENG [1]

The Pennsylvania State University, Department of Electrical Engineering University Park, PA 16802, USA

By generating a blur function as the reference object, the phase of the spectrum of the blur function is automatically canceled in the joint transform spectrum. A reference function is then utilized to compensate the amplitude variation of the blur spectrum in order to restore the original object. Experimental results are given.

1. Introduction

The joint transform correlator (JTC) was proposed by Weaver and Goodman [1] for optically convolving a pair of 2-D functions without using a complex spatial filter in the frequency plane. This idea was further applied to real-time pattern recognition by incorporating some spatial light modulators in the optical system [2]. Recently, deblurring using a nonlinear JTC was suggested [3]. In this paper, we demonstrate experimentally a simple method of deblurring using a joint transform processor [4].

2. Joint transform processor

Shown in fig. 1 is a joint transform processor in which the configuration of the three-lens system between the source point s_1 and the observation plane P is a standard optical information processing system. P_0 is the input object plane and P_1 is the joint transform filter (JTF) plane. Suppose that the amplitude transmittance of the input object is

$$t(x, y) = a_1(x - x_0, y) + a_2(x + x_0, y) . \quad (1)$$

After square law detection at the plane P_1, the amplitude transmittance of the JTF (or joint transform hologram) can be written as

$$t_1(f_x, f_y) = |A_1(f_x, f_y)|^2 + |A_2(f_x, f_y)|^2$$
$$+ 2 \, \mathrm{Re}[A_1(f_x, f_y) A_2^*(f_x, f_y) \exp(-i4\pi x_0 f_x)] , \quad (2)$$

where A_i is the Fourier transform of a_i, $f_x = x/\lambda F$, and $f_y = y/\lambda F$. F is the focal length of all the lenses L, and λ is the wavelength of the light. Re denotes the real part.

In the read-out process, the amplitude transmittance of the reference function $A_3(f_x, f_y)$, in the plane P_2, is imaged by the lens pair L_3 and L_4 onto the JTF through the reflection from the beam splitter BS. Light reflected from the JTF is collected by the lens L_2 and the inverse Fourier transform of $A_3(f_x, f_y) t_1(f_x, f_y)$ occurs at the observation plane P. The output amplitude of interest is either IFT $[A_1(f_x, f_y) A_2^*(f_x, f_y) A_3(f_x, f_y)]$ or IFT$[A_1^*(f_x, f_y) A_2(f_x, f_y) A_3(f_x, f_y)]$, where IFT stands for the inverse Fourier transform. The former is centered at $x = 2x_0, y = 0$ while the latter is centered at $x = -2x_0, y = 0$. In general, the joint transform processor can achieve what a coherent optical processor can do. However, in this paper, we would concentrate on a deblurring problem.

Assume that the amplitude transmittance of the input object takes the form

$$t(x, y) = o(x - x_0, y) * \mathrm{rect}[(x - x_0)/l]$$
$$+ \mathrm{rect}[(x + x_0)/l] \delta(y) , \quad (3)$$

where the first time term is the blurred object. * denotes convolution operation, and l is the blur width.
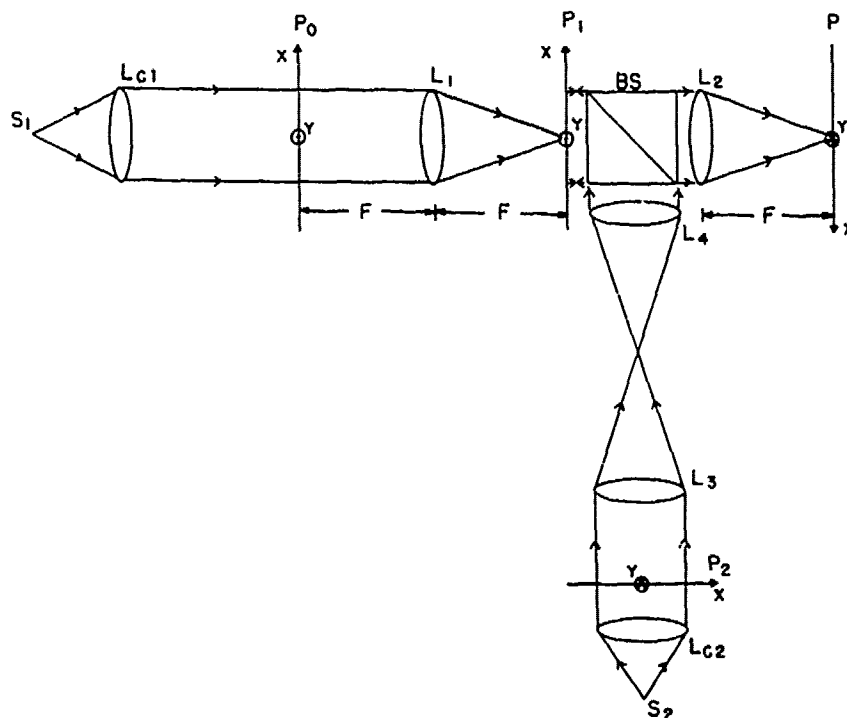
Fig. 1. Joint transform processor.

T^ information in the JTF which is of interest to us is

$$A_1(f_x, f_y) A_2^*(f_x, f_y) = O(f_x, f_y) \, \text{sinc}^2(lf_x) \,. \qquad (4)$$

where the unimportant constants are neglected. We note that the phase of the sinc function is completely canceled due to the prese ·ce of the reference object $\text{rect}(x/l)\delta(y)$. In order to obtain the deblurred image, all we have to do is to generate a reference function $[\text{sinc}^2(lf_x)]^{-1}$ which cancels the undesired amplitude variation $\text{sinc}^2(lf_x)$ in the object spectrum. Thus, the output wave amplitude, which is centered at $x = 2x_0$, $y = 0$, is

$$\text{IFT}[A_1(f_x, f_y) A_2^*(f_x, f_y) A_3(f_x, f_y)] = o(x, y) \,. \qquad (5)$$

If programmable real-time spatial light modulators (such as LCTV are placed in the planes P_0 and P_2, then the references object A_2 and the reference function A_3 can be continuously varied. On the other hand, a real-time square law detector (such as LCLV)

can produce the JTF instantly. The operation $\text{IFT}[A_1 A_2^* A_3]$ can then be achieved in real-time.

3. A deblurring system

Although blurring can occur in different forms, in this paper, we only consider linear blur. In order to reduce the dynamic range requirement, an alternative deblurring system in which Fourier transformation takes place only in one direction is suggested (fig. 2). The recording system and the read-out system can be put together as those in fig. 1 using a beam combiner. However, for the ease of 3-D drawing, we consider them separately. Let the amplitude transmittance of the input plane be

$$t(x, y) = o(x - x_0, y) * \text{rect}[(x - x_0)/l] + \text{rect}[(x + x_0)/l] \,. \qquad (6)$$

Since blurring occurs only in the x-direction, we merely have to perform Fourier transformations in that direction. In the orthogonal direction, a simple
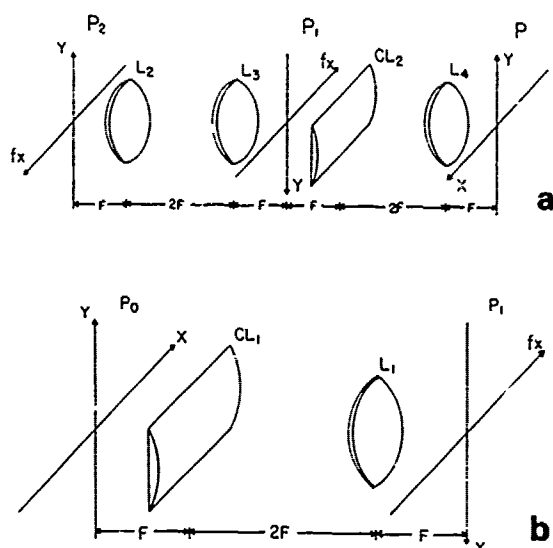
Fig. 2. (a) An optical system for recording of the joint transform spectrum. (b) A read-out optical system.

imaging would suffice. Hence, the information on the blurred object and the reference object are brought together only in the x-direction. After square law detection at the plane P_1, the JTF takes the form

$$t_J(f_x, y) = \text{sinc}^2(lf_x) \; [\,|O(f_x, y)|^2 + 1$$

$$+ O(f_x, y) \exp(-i4\pi x_0 f_x)$$

$$+ O^*(f_x, y) \exp(i4\pi x_0 f_x)\,] , \qquad (7)$$

where $O(f_x, y)$ is the object transformed in the x-direction.

The reference function $[\text{sinc}^2(lf_x)]^{-1}$ on the plane P_2 is then imaged onto the JTF. Finally, an optical system (fig. 2b) similar to that of fig. 2a brings the light to the output plane. Again, inverse Fourier transform is performed in the x-direction while simple imaging is maintained in the y-direction. The wave amplitude at the observation plane P is

$$u(x, y) = o(x, y) * o(x, y) + \delta(x)\,\delta(y)$$

$$+ o(x - 2x_0, y) + o^*(-x - 2x_0, y) . \qquad (8)$$

We note that the third term of the above equation is the deblurred image. For an amplitude object, the fourth term is also a deblurred image.

4. Experiment and discussion

Shown in fig. 3a is the original object in which the characters ABDKU serve as the object. The thin line

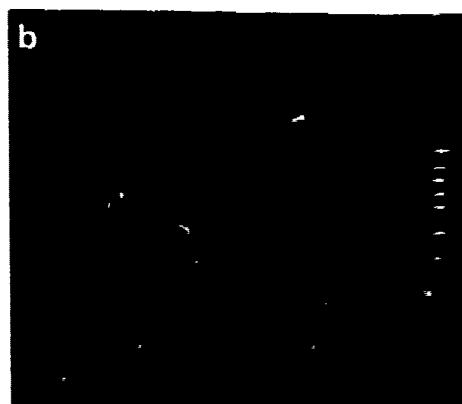Fig. 3. (a) The original object ABDKU with a reference line.

Fig. 3. (b) Blurred object and reference object.

Fig. 3. (c) Joint transform spectrum.

Fig. 3. 'd) Under-exposed JTF to show the fringes in the main-lobe of the spectrum.

Fig. 3. (f) Reference spectrum $1/\text{sinc}^2(lf_x)$.
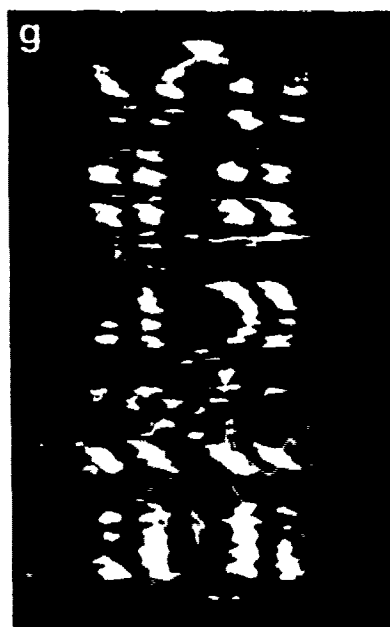
Fig. 3. (g) High-passed deblurred image.

Fig. 3. (e) Overly-exposed JTF to show the fringes in the side-lobes of the spectrum.

at the left hand side of the figure is blurred to be the reference object. By photo reduction, both the characters ABDKU and the reference line are converted to be transparent. The photo-reduced transparency is then placed in the object plane of a simple imaging system. Light from a mercury arc source shining on a piece of ground glass serves a the diffuse light source. Kodak so-253 film is mounted on the camera

above a digital motorized stepper. The motion of the camera causes the recorded image to be blurred. A further contact printing makes the blurred object transparent again (fig. 3b). By comparing fig. 3a with fig. 3b, we note that, since the degree of blur is chosen to be very severe, the blurred image is beyond recognition. The distance between the blurred object and the reference object is about 6 mm. We have chosen the original line to have a finite width so that the reference object can have comparable transmittance as that of the blurred object. In the above re-

22

cording processes. linearity is ensured by pre-exposing the film to a uniform light with total exposure somewhat above the toe of the H&D curve.

In the system of JTF recording, the focal length of the cylindrical lens CL_1 is 30 cm and that of the spherical lens L_1 is 20 cm. Their separation is approximately the sum of their focal lengths. Hence, the image is compressed in the y-direction by a factor 2/3. Shown in fig. 3c is the joint transform spectrum recorded on the Kodak 649-F plate. Because we have exposed the blurred object in the linear region, the dark region of fig. 3b is not totally opaque and the average transmitted light is squeezed into a vertical line by the lens L_1 which overly exposes the zero-frequency spectrum. Some regions in the spectrum are also overly exposed so that the interference fringes are quite lost. To reduce the effect of this problem. more light from the reference object is required in order to faithfully record the spectrum. This can be done by placing a filter in front of the characters ABDKU during the blurring process. In fig. 3d, the film is under-exposed to show the fringes in the mainlobe of the spectrum. On the other hand, in fig. 3e, the film is overly exposed to reveal the fringes in the sidelobes. The reference function $[sinc^2(l/f_r)]^{-1}$ can best be generated by computer. However, in the experiment, an approximate reference function is produced by exposing a 649-F plate to the spectrum of the reference object only. The negative of the result is shown in fig. 3f in which the nonuniformity in the vertical direction is caused by that in the reference object.

During the read-out process, the reference function is imaged by a lens pair L_2 and L_3, each of focal length 30 cm, onto the JTF (fig. 3c) on the plane P_1. The focal length of both the cylindrical lens CL_2 and the spherical lens L_4 is 30 cm. Again, they are separated by the sum of their focal lengths. Many reference functions and joint transform filters with different exposures and different chemical processing conditions are experimented. The image of the original object which shows some degree of deblurring is shown in fig. 3g. We note that the image is stretched in the horizontal direction by a factor of 9/4 with respect to the vertical direction because of the particular lenses used. Owing to the presence of the vertical lines in both the joint transform filter and the spectrum of the reference function. a high-passed

image is expected. However, due to the reason that is not quite clear, the edges of the images are not quite sharp. The horizontal parts of all the characters are quite lost because of overexposure of the corresponding regions which result in the lost of the interference fringes. This problem would be avoided if a continuous tone object is utilized as the original object. If the image of the original object instead of the high-passed object is expected, the unwanted dark region of fig. 3b should be painted totally opaque before performing the read-out process.

Deblurring is mainly a combined process of phase correction and high-pass filtering. Since the phase cancellation is automatically achieved in the JTF, we merely have to perform a high-pass filtering process. A simpler experiment is also performed in which the mainlobe of the joint transform spectrum is overly exposed to generate the approximate high-pass function directly on the JTF. The read-out optical system would be a cylindrical lens together with a spherical lens. Comparable result as fig. 3g is obtained.

5. Conclusion

A simple deblurring method using a joint transform processor is suggested. By incorporating some real-time devices into the optical system. the process can potentially be carried out in real-time. Although the experiment is performed with photographic films, the result nevertheless shows some degree of deblurring. If a continuous tone object is used instead and the procedures discussed in section 4 are taken into account. the deblurred image would be much better. Some other methods. such as the one suggested by Yang and Leith [5], can also be utilized to restore the original object.

Based on the simple method we have suggested. for a general blur (defocusing, nonlinear motion, etc.), we would generate the same blur as the reference object. The phase of the blur spectrum is automatically canceled in the joint transform spectrum. Then, we would generate a reference function to cancel the nonuniformity of the blur spectrum. Thus, a deblurred image can be obtained.

Acknowledgement

References

[1] C.S. Weaver and J.W. Goodman, Appl. Optics 5 (1966) 1248.
[2] F.T.S. Yu and X J. Lu, Optics Comm. 52 (1984) 10.
[3] B. Javadi, H.J. Caulfield and J.L. Horner, Technical Digest, Optical Society of America Annual Meeting (1988) 89.
[4] F.T.S. Yu and J.E. Ludman, Micro Opt. Tech. Lett. 1 (1988) 374.
[5] G.G. Yang and E. Leith, Optics Comm. 36 (1981) 101.

RESTORATION OF BLURRED IMAGES DUE TO LINEAR MOTION USING A JOINT TRANSFORM PROCESSOR

Feng Cheng, Xin Xu, Shudong Wu, and Francis T. S. Yu
Department of Electrical Engineering
The Pennsylvania State University
University Park Pennsylvania 16802

Don A. Gregory
U.S. Army Missile Command
Redstone Arsenal Alabama 35898

KEY TERMS

Image processing, joint transform processor, Fourier transform

ABSTRACT

*In this paper we shall show that blurred photographic images due to
linear motion can be restored with a joint transform processor (JTP)
The deconvolution function for the image deblurring can be directly
implemented at the input plane of a JTP. The bipolarity of the deconvo-
lution function can be accomplished by encoding with a π-shifted grating
structure. Computer simulations to test the feasibility of the scheme are
given.*

One of the major advantages of a joint transform architecture
[1, 2] must be the avoidance of the filter synthesis, and the
architecture of the processor can also be applied to generai-
ized image processing [3]. Research in this area has been
motivated by recent advances in spatial light modulator tech-
nology and the inherent real-time processing capability.

This paper demonstrates the implementation of a joint
transform processor (JTP) architecture to restore blurred im-
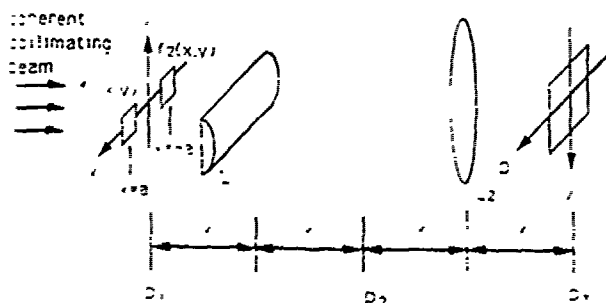
Figure 1 A one-dimensional joint transform processor

ages due to linear motion. In a conventional coherent optical processor. image deblurring can be performed by synthesizing a restoration filter. whose transfer function is the inverse of the blurred function [4]. Notice that the inverse filter is generally not physically realizable due to many concomitant difficulties. In a JTP. however. the object is made to convolve with a processing (reference) function that is placed at the input plane. In this arrangement. some of the processing constraints may be alleviated. Because the deblurring function is generally a bipolar function. we have proposed a spatially encoding technique to produce the bipolarity of the processing function. In other words. two sets of grating of the same frequency are encoded on the input transparency to produce a π phase shifted vector. such that a bipolar nature of the processing function can be obtained in the first-order diffraction.

Figure 1 shows a 1-D joint Fourier transform correlator. where cylindrical lens L_1 has no effect in the x direction. Thus. the Fourier transforms of the object and the reference function in this direction are obtained at P_1 by L_2. In the y direction. however. the input functions are imaged onto P_3 by the combination of the two lenses L_1 and L_2. Thus. a 1-D joint transform hologram can be obtained at P_3 with a square-law detector (SLD). We assume that an object transparency $f_1(x, y)$ and a reference $f_2(x, y)$ are inserted in the

input plane. The transmittance distribution can be written as 87

$$t(x, y) = f_1(x - a, y) - f_2(x + a, y) \qquad (1)$$

where $2a$ is the separation of f_1 and f_2. By coherent illumination. the complex light distribution at the Fourier plane is given by

$$T(p) = F_1(p)\exp(-iap) - F_2(p)\exp(iap) \qquad (2)$$

where p represents the angular spatial frequency coordinate.

$$F_1(p) = F\{f_1(x, y)\}, \qquad F_2(p) = F\{f_2(x, y)\} \quad (3)$$

and F denotes a 1-D Fourier operator. The transmittance of the joint transform hologram produced by a square-law detector can be written as

$$|T(p)|^2 = |F_1|^2 + |F_2|^2 - F_1 F_2^* \exp(-2ap) - F_1^* F_2 \exp(2ap) \qquad (4)$$

The system setup for the readout process is the same as shown in Figure 1. By placing the joint transform hologram at the input plane. the cross correlation of functions f_1 and f_2 can be observed at $(2a, 0)$ and $(-2a, 0)$ of the output plane. If one of the two functions is flipped along the x direction. for example. $f_1(x - a, y)$. then the convolution operation instead of the correlation can be performed. It is therefore evident that JTP can be used as a generalized image processor.

Let us assume that the blurred function is

$$g(x) = o(x) * \text{rect}\left(\frac{x}{L}\right) \qquad (5)$$

where $o(x)$ represents the unblurred object.

$$\text{rect}\left(\frac{x}{L}\right) = \begin{cases} 1, & -L/2 \leq x \leq L/2 \\ 0, & \text{otherwise} \end{cases}$$

is the linear smearing function. L is the smearing length. and * denotes the convolution operation. It is apparent that. for general image processing. the output complex light field can
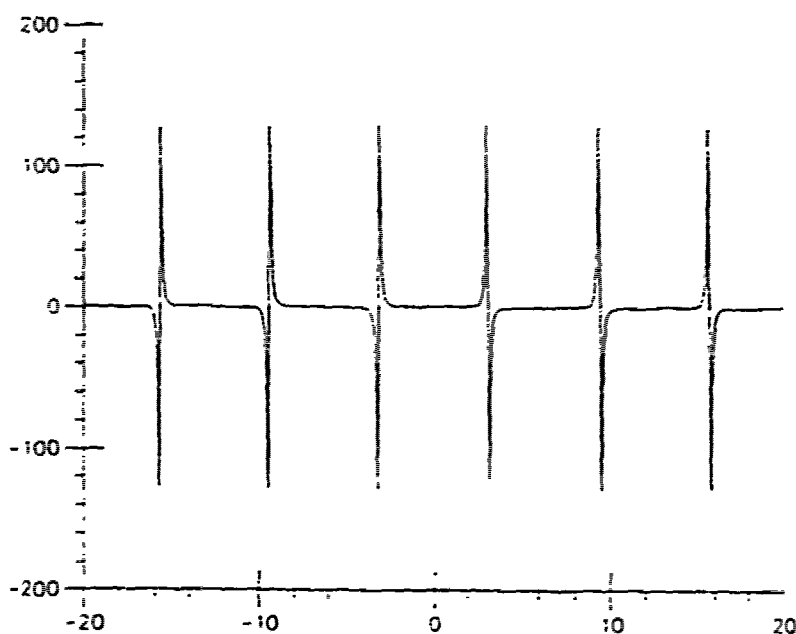

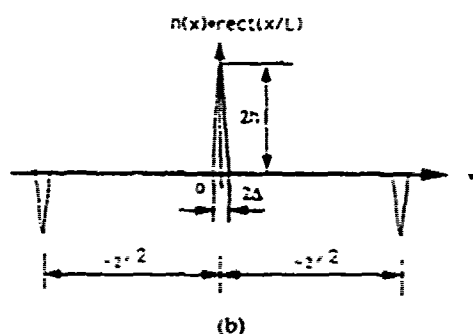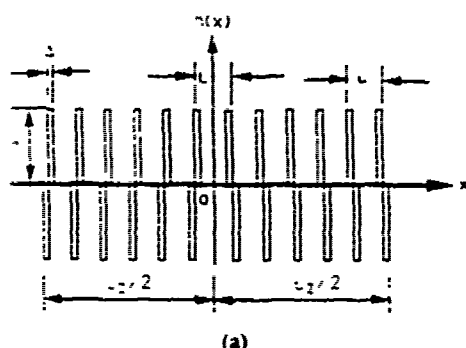**Figure 2** The shape of the processing function. given $o = 1$ ($L = 2$), $a = 0$)

(a)

Figure 4 Grating encoding pattern applied to the processing function

(b)

Figure 3 (a) An example of modified processing function; (b) convolution of (a) with the smearing function

(a)   (b)
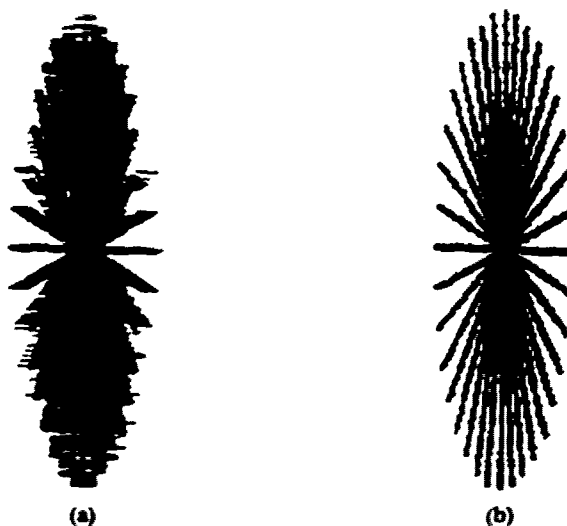
Figure 5 (a) A blurred image due to linear motion; (b) output deblurred image from (a)

be written as

$$r(x) = g(x) * h(x) = o(x) * \text{rect}\left(\frac{x}{L}\right) * h(x) \quad (6)$$

where $h(x)$ is the spatial impulse response of the filter function. It can be shown that, for a linear motion image deblurring, the spatial impulse response of a deblurring filter is [see the appendix]

$$h(x) = k \frac{\sin(\omega|x|)\sinh(\omega a)}{[\cos(\omega x) - \cosh(\omega a)]^2} \quad (7)$$

where k is a proportionality constant and a is an arbitrarily small positive parameter that determines the width of the spikes of $h(x)$ as shown in Figure 2. However, in practice, the deblurring function $h(x)$ is spatially limited, and the spikes are reshaped into rectangular form for optical implementation, as used by Swindell [5] and shown in Figure 3(a). The corresponding restored point image is depicted in Figure 3(b). To separate the image from the ghost images, it is required that the size of the reference should be larger than double the object size, i.e., $L_2 > 2L_1$.

The bipolarity of the deblurring function $h(x)$ can be obtained by encoding a set of π-shifted Ronchi gratings [6], as shown in Figure 4. The transmittance of the encoded transparency can be written as

$$T_1(x, y) = h_+(x) \sum_{k=-\infty}^{\infty} \delta(y - 2bk)\text{rect}(y/b)$$

$$- h_-(x) \sum_{k=-\infty}^{\infty} \delta(y - 2bk - b)\text{rect}(y/b) \quad (8)$$

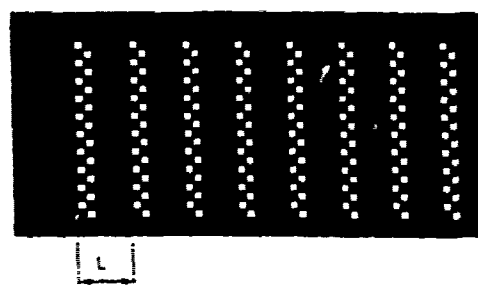where $2b$ is the grating period, and $h_+(x)$ and $h_-(x)$ are the positive and negative functions of $h(x)$. By 1-D Fourier transforming the above equation with respect to the y axis, it becomes

$$T_1(x, q) = \sum_{k=-\infty}^{\infty} \delta(q - k\pi/b)[h_+(x) - \exp(jbq)h_-(x)] \quad (9)$$

where q is the spatial frequency coordinate. The first-order deblurred image can be obtained from the joint transform hologram as expressed in Equation (6). It is therefore apparent that the image deblurring can be obtained with the proposed 1-D JTP.

The feasibility of the optical implementation is studied by a computer simulation as given in Figure 5. Figure 5(a) shows a blurred image with a smearing length $L = 10$ pixels. The rectangular spikes of the deblurring function $h(x)$ are 1 pixel wide. The corresponding computer-simulated result is shown in Figure 5(b).

Finally, we would conclude that restoring blurred images using a one-dimensional joint transform processor is feasible. The bipolarity of the processing function can be alleviated with a grating encoding scheme. This system, in principle, can be implemented with real-time devices.
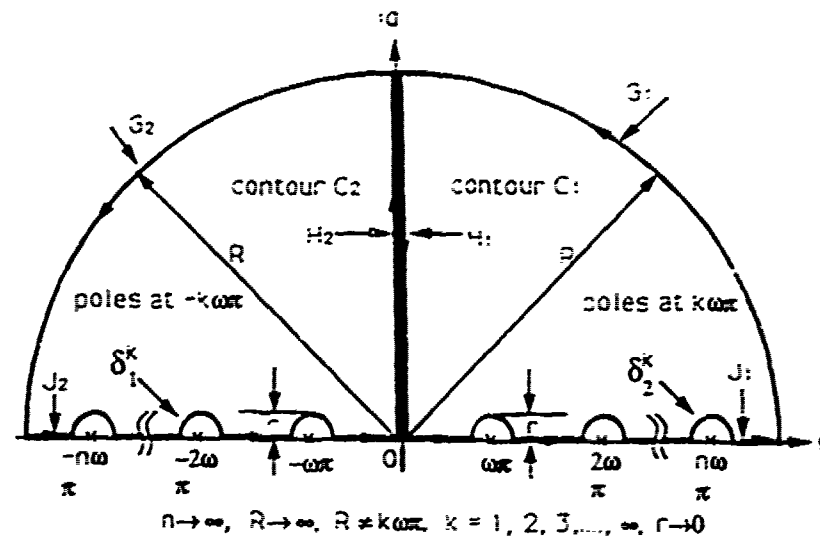
Figure 6 The contours of the complex integrals for calculating the processing function

APPENDIX

The spatial impulse response function $h(x)$ can be expressed by the Cauchy principle value with the inverse Fourier transform of an inverse filter. such as

$$h(x) = \int_{-\infty}^{-\infty} \exp(-\alpha|p|) \left[\mathrm{Sinc}\left(\frac{p}{\omega}\right) \right]^{-1} \exp(jxp)\, dp$$

$$= 2\int_{0}^{-\infty} \exp(-\alpha p) \left[\mathrm{Sinc}\left(\frac{p}{\omega}\right) \right]^{-1} \cos(xp)\, dp \quad (10)$$

where $\alpha \to 0$. A closed form of $h(x)$ can be derived. based on the sum of two contour integrations. given by

$$I = \int_{C_1} \exp(-\alpha z) \left[\mathrm{Sinc}\left(\frac{z}{\omega}\right) \right]^{-1} \exp(jxz)\, dz$$

$$- \int_{C_2} \exp(\alpha z) \left[\mathrm{Sinc}\left(\frac{z}{\omega}\right) \right]^{-1} \exp(jxz)\, dz \quad (11)$$

where contours C_1 and C_2 are shown in Figure 6. When $x \geq 0$. $I = 0$ and it can be shown that

$$h(x) = -2\pi^2 \sum_{k=1}^{\infty} (-1)^k \exp(-\alpha k\omega)\sin(k\omega x)$$

$$= 2\pi^2 \frac{\sin(\omega x)\sinh(\omega\alpha)}{\left[\cos(\omega x) + \cosh(\omega\alpha)\right]^2} \quad (12)$$

When $x < 0$. $h(x) = h(|x|)$ because $h(x)$ is an even function. Thus. the general form of $h(x)$. as in Equation (7). can be obtained.

ACKNOWLEDGMENT

REFERENCES

1. C. S. Weaver and J. W. Goodman. "A Technique for Optically Convolving Two Functions." *Appl. Opt.*, 5. 1966. p. 1248.
2. F. T. S. Yu and J. E. Ludman. "Microcomputer-Based Programmable Optical Correlator for Automatic Pattern Recognition." *Opt. Lett.*, 11. 1986. p. 395.
3. F. T. S. Yu and J. E. Ludman. "Joint Fourier Transform Processor." *Microwave Opt. Tech. Lett.*, 1. 1988. p. 374.
4. F. T. S. Yu. *Optical Information Processing*. John Wiley & Sons. New York. 1983. Chapter 7.
5. W. Swindell. "A Noncoherent Optical Analog Image Processor." *Appl. Opt.*, 9. 1970. p. 2459.
6. G. G. Yang and E. Leith. "An Image Deblurring Method Using Diffraction Gratings." *Opt. Commun.*, 36. 1981. p. 101.

APPENDIX 10.11

Reconfigurable Interconnections

Reconfigurable interconnections using photorefractive holograms

Shudong Wu, Qiwang Song, Andy Mayers, Don A. Gregory and Francis T. S. Yu

Using coupled wave theory and the law of refraction. diffraction properties of volume holograms are discussed. Reconfigurable interconnections by either wavelength tuning or spatial division techniques are proposed. Reflection type volume holograms can be used for a large number of reconfigurable interconnections in terms of finite wavelength tunability. Transmission volume holograms encoded in pinhole holograms can be easily reconfigured by spatial light modulator. Experimental demonstrations obtained by using these methods are presen*ed.

I. Introduction

In recent years, high speed communicat.. n and high speed computation have aroused interest in optical interconnections.[1] The advantages of optical interconnections over electrical interconnections are high speed, massive parallelism, and lower mutual interference. Optical fibers, integrated waveguides, and free-space interconnections are all being considered. However, free-space interconnection is the most attractive among the others for its potential massive parallelism. Both classical and holographic optical elements are being considered as candidates for free-space interconnections. Generally, classical optical elements have higher efficiency of light utilization; nevertheless the holographic optical elements have greater flexibility.

Reconfigurable massive interconnection is important for communication and computation aspects. Spatial light modulators (SLMs) were proposed for use as a reconfigurable optical crossbar[2] based on optical matrix-vector multiplication. The main problem with the SLM method is low energy efficiency. In addition, the SLMs suffer from low contrast, low resolution. slow response, and large physical size. To increase the energy efficiency, Yeh et al.[3] proposed a two-wave mixing technique to amplify the controlling pattern, which is generated by the SLM. However, high light intensity (~ 1 W/cm^2) is required for high

Qiwang Song is with the Department of Electrical and Computer Engineering. Syracuse University. Syracuse. New York 13244. Don Gregory is with the U S. Army Missile Command. Redstone Arsenal. Alabama 35898. and the other authors are with Pennsylvania State University. Department of Electrical Engineering. University Park. Pennsylvania 16802.

efficiency of energy transference, and the method may not be applicable at the present time.

In recent years, laser technology has rapidly advanced with compact wavelength tunable laser diodes becoming available. Their wavelength tunability provides a convenient method of controlling the direction of diffracted light beams. Lin[4] has proposed a technique for using a multiplexed waveguide transmission hologram and wavelength tunable light source to realize the reconfigurable optical interconnections. Generally speaking, the wavelength tunable range $\Delta\lambda/\lambda$ of laser diodes is $\sim 10^{-2}$.[5] However, for transmission holograms. 10^{-2} relative wavelength variation would produce a very small deflection of the laser beam. On the other hand, with a surface waveguide, the usable space is essentially one-dimensional. These are the major reasons that limit the number of interconnections using transmission holograms and surface waveguides. Multifacet computer-generated holograms may also be used for reconfigurable optical interconnections. However, for high fanout interconnections, the energy efficiency of the holograms is generally low

It is well known that a photorefractive volume hologram can have high diffraction efficiency and high angular and wavelength selectivities. We first discuss the performance of transmission and reflection holograms based on coupled wave theory and law of refraction. Then reconfigurable optical interconnection schemes using wavelength tuning and spatial division techniques are proposed. A small amount of wavelength tuning can be used to realize a large number of reconfigurable interconnections. which can be obtained by choosing the writing angle and thickness of the hologram. For the spatial division method. the interconnection patterns can be recorded as a series of pinhole holograms in the photorefractive material.

This series of pinhole holograms can be reconstructed with an array of light sources and by simply translating a pinhole generated by an SLM, and the interconnection pattern can be reconfigured. We shall experimentally demonstrate the reconfigurabilities of both techniques using a nonlinear photorefractive crystal, LiNbO₃.

II. Properties of Volume Holograms

We first discuss some basic properties of volume holograms. We assume that two write-in beams forming an angle 2θ are entering into a photorefractive material as shown in Fig. 1. The vector magnitude of the interference grating within the crystal can be written as

$$|K| = \frac{2\pi \sin\theta}{\lambda} . \tag{1}$$

where λ is the writing wavelength and n is the average refractive index of the crystal. If the readout wavelength is deviated from λ by $\Delta\lambda$, the Bragg diffraction condition requires a change of $\Delta\theta$. By differentiating Eq. (1) with respect to λ and θ, respectively, the relationship between $\Delta\lambda$ and $\Delta\theta$ is given by

$$\frac{\Delta\lambda}{\lambda} = \cot(\theta)\Delta\theta - \frac{1}{2}\Delta\theta^2 . \tag{2}$$

for which the change of λ should be compensated by the change of θ. We call 2θ the internal writing angle within the crystal. Note that the external writing angle 2α is different from the internal writing angle 2θ. By applying the law of refraction to Eq. (2) for the unslanted transmission holograms in Fig. 1(a), we have

$$\Delta\theta = \frac{\cos\alpha}{\sqrt{n^2 - \sin^2\alpha}} \Delta\alpha. \tag{3}$$

$$\left\{\frac{\Delta\lambda}{\lambda}\right\} = \cot(\alpha)\Delta\alpha - \frac{1}{2}\frac{\cos^2\alpha}{n^2 - \sin^2\alpha}(\Delta\alpha)^2. \tag{4}$$

Similarly as for the unslanted reflection hologram in Fig. 1(b), we have

$$\Delta\theta = \frac{\sin\alpha}{\sqrt{n^2 - \cos^2\alpha}} \Delta\alpha. \tag{5}$$

$$\left\{\frac{\Delta\lambda}{\lambda}\right\} = \frac{\sin\alpha \cos\alpha}{n^2 - \cos^2\alpha}\Delta\alpha - \frac{1}{2}\frac{\sin^2\alpha}{n^2 - \cos^2\alpha}(\Delta\alpha)^2. \tag{6}$$

where 2α is the external write-in angle. We note that only the second-order term in Eq. (6) is significant as α approaches 90°. However, we shall consider cases for α in the 10–80° region, and all second-order terms can be neglected. The reciprocal of the dispersions

$$\left(\text{i.e.,} \frac{1}{\lambda}\frac{d\lambda}{d\alpha} \right)$$

for transmission and reflection holograms are functions of external write-in angle α in Eqs. (4) and (6) and are shown in Figs. 2(a) and (b). We see that under the same write-in angle 2α, the reflection hologram would provide higher dispersion, which is due to the refraction at the surface of the photorefractive crystal. Since the refractive index of the crystal is higher than
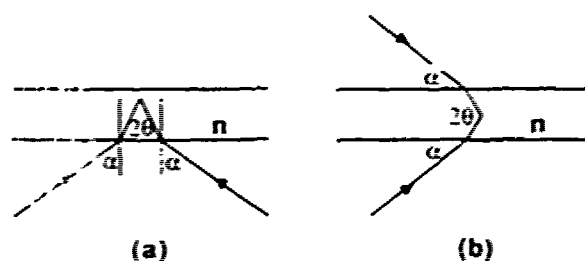

(a)   (b)

Fig. 1   Writing angle inside and outside the recording media: (a) transmission hologram; (b) reflection hologram.
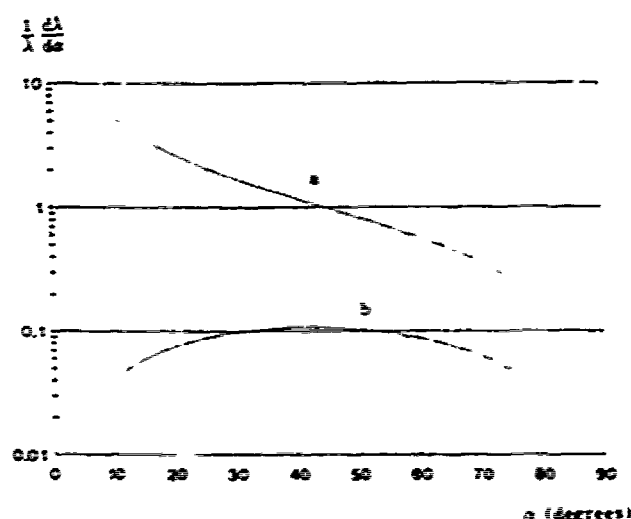

Fig. 2   Dispersion $(1/\lambda)$ $(d\lambda/d\alpha)$ as a function of half-writing angle α. $n = 2.25$: a, transmission hologram; b, reflection hologram.

the index of air, the internal angle θ is less than the external angle α for the transmission hologram and $\theta > \alpha$ for the reflection hologram. This means that for the same external angle α, the reflection hologram would have higher dispersion. On the other hand, when α approaches 0°, the angular magnification $(\Delta\alpha/\Delta\theta)$ becomes larger due to refraction, as can be deduced from Eq. (5). This would also cause higher dispersion in reflection holograms. In the region between $\alpha = 30$ and 60°, a uniform dispersion is observed for the reflection hologram.

To analyze the angular and wavelength selectivities of volume holograms, we shall refer to the coupled wave theory obtained by Kogelnik.[6] Let us assume that the multiple diffraction process can be neglected for a multiexposure hologram. The problem can be treated with the weak coupling approximation. Thus, by referring to the coupled wave theory, the normalized diffraction efficiency for an unslanted transmission hologram can be expressed as

$$\tau = \frac{\sin^2(\nu^2 + \xi^2)^{1/2}}{(1 + \xi^2/\nu^2)} . \tag{7}$$

where $\nu = \pi n_1 d/\lambda \cos\theta$. n_1 is the amplitude of the spatial modulation of the refractive index. d is the thickness of the hologram.

$$\zeta = \frac{2\pi n d \sin\theta}{\lambda} \Delta\theta$$

$$= -\frac{2\pi n d}{\lambda} \frac{\sin^2\theta}{\cos\theta} \frac{\Delta\lambda}{\lambda}$$

is known as the dephasing parameter, which is dependent on the reading angular deviation $\Delta\theta$ and the reading wavelength deviation $(\Delta\lambda/\lambda)$.

For the weak coupling approximation (i.e., $|\nu| \ll |\zeta|$), Eq. (7) can be simplified, as given by

$$\eta = \nu^2 \text{sinc}^2\zeta. \tag{8}$$

The angular and wavelength selectivities of the holograms can be defined by the full width of the main lobe of the sinc function and are given by

$$\{\Delta\theta\}_t = \frac{\lambda}{nd\sin\theta}, \tag{9}$$

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_t = \frac{\lambda\cos\theta}{nd\sin^2\theta}, \tag{10}$$

where subscript t signifies transmission hologram.

With reference to Eq. (3) and the law of refraction, we find that

$$\{\Delta\alpha\}_t = \frac{\sqrt{n^2 - \sin^2\alpha}}{\sin\alpha\cos\alpha} \frac{\lambda}{d}, \tag{11}$$

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_t = \frac{\sqrt{n^2 - \sin^2\alpha}}{\sin^2\alpha} \frac{\lambda}{d}. \tag{12}$$

Similarly, for an unslanted reflection hologram, the normalized diffraction efficiency can be expressed as

$$\eta = \frac{1}{\left[1 + \left\{1 - \frac{\zeta^2}{\nu^2}\right\} / \text{sh}(\nu^2 - \zeta^2)^{1/2}\right]}, \tag{13}$$

where

$$\nu = j\pi n_1 d/\lambda \sin\theta,$$

$$\zeta = \frac{2\pi n d \cos\theta}{\lambda} \Delta\theta$$

$$= \frac{-2\pi n d \sin\theta}{\lambda} \frac{\Delta\lambda}{\lambda}.$$

For the weak coupling case (i.e., $|\nu| \ll |\zeta|$), Eq. (13) can be reduced to the following expression:

$$\eta = |\nu|^2 \text{sinc}^2\zeta. \tag{14}$$

The equations for angular and wavelength selectivities of an unslanted reflection hologram can be derived, such as

$$\{\Delta\alpha\}_r = \frac{\sqrt{n^2 - \cos^2\alpha}}{\sin\alpha\cos\alpha} \frac{\lambda}{d}, \tag{15}$$

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_r = \frac{1}{\sqrt{n^2 - \cos^2\alpha}} \frac{\lambda}{d}, \tag{16}$$

where subscript r signifies the reflection hologram.

Figure 3 shows the angular selectivities, $(\Delta\alpha)_t$ and $(\Delta\alpha)_r$, as functions of an external write-in angle α, where the values of $\Delta\alpha$ have been normalized by λ/d.
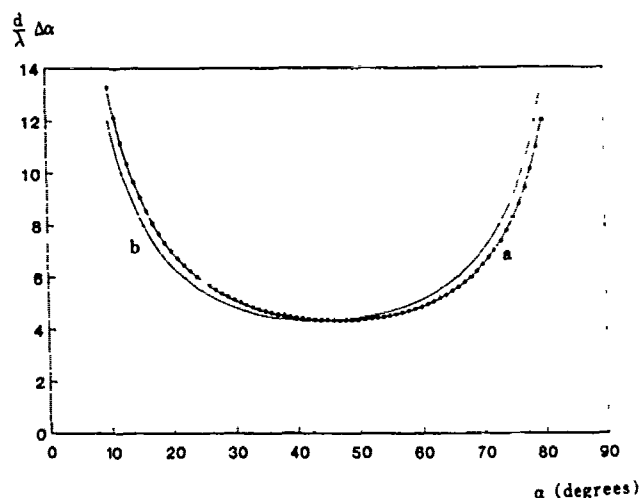

Fig. 3. Angular selectivity of volume holograms normalized by λ/d, $n = 2.28$: a, transmission hologram; b, reflection hologram.
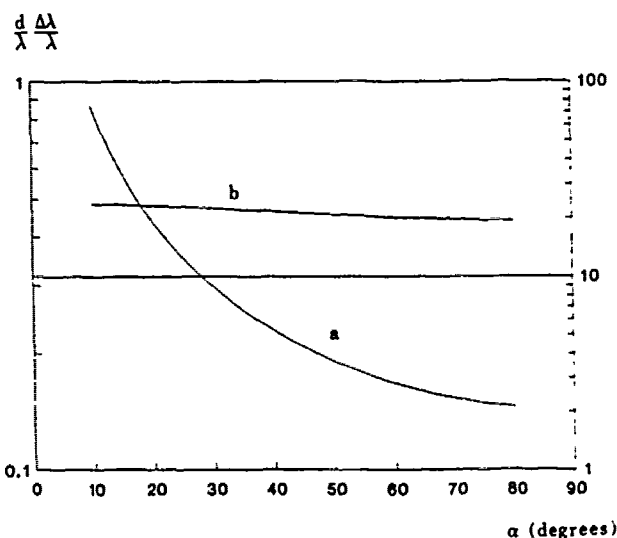

Fig. 4. Wavelength selectivity of volume holograms normalized by λ/d, $n = 2.28$: a, transmission hologram (the scale on right); b, reflection hologram (the scale on the left).

From this figure we see that the optimal values occur at

$$(\alpha_t)_{\text{opt}} = \arcsin\sqrt{n(n - \sqrt{n^2 - 1})}, \tag{17}$$

$$(\alpha_r)_{\text{opt}} = \arccos\sqrt{n(n - \sqrt{n^2 - 1})}, \tag{18}$$

respectively. For example, if $n = 2.28$ for LiNbO$_3$ crystal, $(\alpha_t)_{\text{opt}} = 46.5°$ and $(\alpha_r)_{\text{opt}} = 43.5°$. We note further that in the region of $\alpha = 30$–$60°$, the hologram has the sharpest angular selectivity both for transmission and reflection holograms.

Figure 4 shows the wavelength selectivities

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_t \text{ and } \left\{\frac{\Delta\lambda}{\lambda}\right\}_r$$

as functions of α. From this figure we see that the wavelength selectivity for the transmission hologram
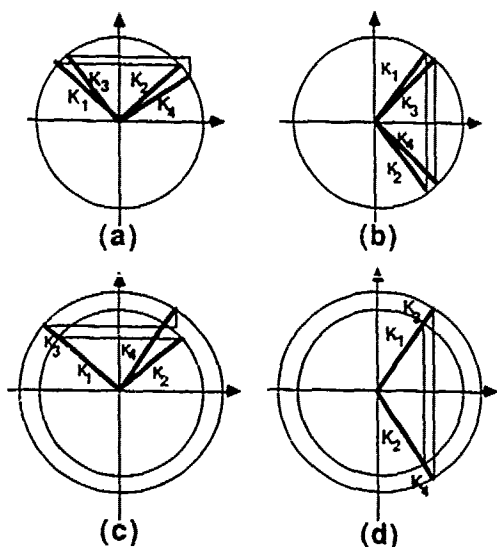
Fig. 5. Diffracted wave vectors derived according to momentum conservation (the horizontal components of the dephasing vector should always be zero): (a) transmission hologram with incident angular deviation; (b) reflection hologram with incident angular deviation; (c) transmission hologram with reading wavelength deviation; (d) reflection holoram with reading wavelength deviation.
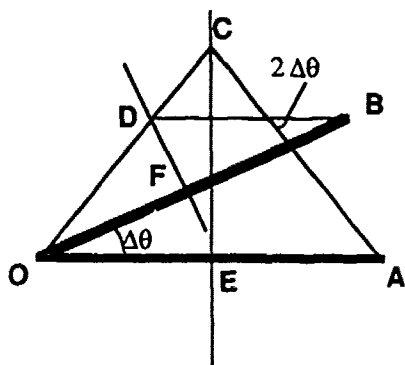

Fig. 6. Principle of reconfigurable optical interconnects with volume holograms. OA and OB are two grating vectors with the same magnitude and an angular separation $\Delta\Theta$. CE and DF are the perpendicular bisectors of OA and OB.

increases as α increases. However, transmission holograms generally have a poorer wavelength selectivity than the reflection hologram. The wavelength selectivity of the reflection hologram is relatively unform over α, which is $\sim 0.45\lambda/d$.

With reference to the principle of momentum conservation, the diffracted wave vectors can be determined. Let us assume that the holographic recording medium is infinitely extended in the x- and y-directions and has a finite thickness d in the z-direction. A dephasing wave vector can be defined as

$$\Delta\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3 + \mathbf{k}_4, \tag{19}$$

where $\mathbf{k}_1$ and $\mathbf{k}_2$ are the two writing wave vectors, $\mathbf{k}_3$ is the reading wave vector, and $\mathbf{k}_4$ is the diffracted wave

vector. In terms of momentum conservation, it is required that

$$(\Delta k)_x = (\Delta k)_y = 0. \tag{20}$$

Thus the z-component of Δk is directly related to the dephasing parameter ζ in Eq. (7) by

$$\zeta = (\Delta k)_z \cdot d. \tag{21}$$

Equation (7) can be used to determine the diffracted wave vector as shown in Fig. 5. For simplicity, we have assumed that all the wave vectors are in the x-z plane. Figure 5(a) shows that the reading angle of the transmission hologram deviates the amount $\Delta\Theta$ from the write-in angle Θ. By referring to Eq. (20), the diffracted wave vector should deviate $-\Delta\Theta$ in the first-order approximation. The wave vector diagram for the reflection hologram is shown in Fig. 5(b). In this figure we see that the angular deviations of the reading and writing wave vectors are identical. Figures 5(c) and (d) show the effects due to reading wavelength dephasing. For the transmission hologram [i.e., Fig. 5(c)], the diffracted wave vector deviates by $\Delta\Theta$; for the first-order approximation, we have

$$\Delta\Theta = 2\tan\Theta\,\frac{\Delta\lambda}{\lambda}. \tag{22}$$

This indicates that for a finite wavelength bandwidth, it will cause the diffracted focal spot to smear. However, reflection holograms do not suffer from such focal spot smearing, since the wave vectors for unslanted reflection holograms are always symmetrical about the x-axis.

III. Reconfigurability with Wavelength Tuning

The principle of the reconfigurable interconnection using wavelength tuning can be seen as illustrated in Fig. 6. The grating vectors OA and OB are written in the photorefractive crystal with an angular separation $\Delta\Theta$. When the grating OA is read out with wave vector OC, the diffracted wave vector is denoted by AC (CE is the perpendicular bisector of OA). On the other hand, if grating OB is read out in the same direction, the magnitude of the reading vector should be changed to OD (DF is the perpendicular bisector of OB), and the diffracted wave vector deviates from AC by $2\Delta\Theta$. Thus, by changing the wavelength, we can read different gratings in the same reading direction. Since different gratings cause different diffraction angles, the reconfigurable optical interconnections can be realized in this simple fashion. The interconnection pattern is previously recorded in the volume hologram, for which the angular and wavelength selectivities should be high enough to prevent crosstalk between channels. We assume that the range of tunable wavelengths

$$\left[\frac{\Delta\lambda}{\lambda}\right]$$

of a laser diode is ~ 0.01 (Ref. 5) (i.e., $\Delta\lambda/\lambda \approx 0.01$), and the required angular deviation is $2\Delta\alpha = 0.2$ rad, then α should be $>84°$ for transmission holograms [from Eq.

(2)]. We note that this angle requires a large size recording crystal (i.e., about $a/\cos\alpha$, where a is the laser beam diameter). However, for reflection holograms, α is $\sim 45°$ [from Fig. 2(b)], which is in fact the optimum angle for angular selectivity, and this angle is also more convenient to work with in practice.

The choice of thickness d of the photorefractive crystal should allow the bandwidth of the hologram to be equal to or greater than the signal bandwidth. If the modulation frequency of the laser diode is Δf,

$$\left\{\frac{\Delta\lambda}{\lambda}\right\} = \frac{\lambda\Delta f}{c},$$

where c is the velocity of light and subscript s signifies the signal. By substituting this relationship in Eq. (16), we see that the thickness d of the crystal should satisfy the following inequality:

$$d \le \frac{\lambda}{\sqrt{n^2 - \cos^2\alpha}}\frac{1}{\left\{\frac{\Delta\lambda}{\lambda}\right\}_s} = d_{max}. \qquad (23)$$

Notice that the corresponding angular selectivity of $(\Delta\alpha)_r$ can be obtained from Eq. (15). By taking $(\Delta\alpha)_c = 3(\Delta\alpha)_r$ as the angular separation between channels, the crosstalk due to the sidelobes of the sinc function would be less than -20 dB, where the subscript c signifies the channel. The wavelength interval

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_c$$

can be derived from Eq. (6), and it can be used as the longitudinal mode spacing of laser diodes for wavelength tuning. This would require the effective optical cavity length to be

$$n_l L = \frac{\lambda}{\frac{c}{\frac{\Gamma}{c}}\left[\frac{\Delta\lambda}{\lambda}\right]_c}, \qquad (24)$$

where n_l is the refractive index of the laser medium, and L is the physical length of the cavity.

As an example using LiNbO$_3$ crystal ($n_0 = 2.28$) as the recording medium, $\alpha = 45°$, $\Delta f = 10$ GHz, and $\lambda = 680$ nm, the relative signal bandwidth is

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_s = 2.4 \times 10^{-5},$$

and the thickness of the photorefractive crystal $d_{max} = 13.8$ mm, the angular selectivity is $(\Delta\alpha)_r = 2.1 \times 10^{-4}$, and the angular separation between channels is $(\Delta\alpha)_c = 6.3 \times 10^{-4}$. For 0.1-rad angular deviation, one would have about 160 channels with a wavelength tuning range $\Delta\lambda/\lambda$ of ~ 0.01. The wavelength interval between channels can be shown as

$$\left\{\frac{\Delta\lambda}{\lambda}\right\}_c = 6.7 \times 10^{-5},$$

and the effective cavity optical length of the laser diode is $n_l L = 5$ mm. If

$$0.1\left\{\frac{\Delta\lambda}{\lambda}\right\}_s$$

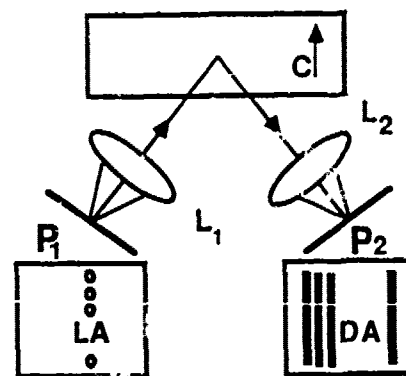is used as the required relative wavelength stability, we would have the required wavelength stability of laser

Fig. 7. Reconfigurable optical crossbar: C, z-cut LiNbO$_3$ crystal; L_1, L_2, collimating and focusing lens; P_1, P_2, input and output planes; LA, DA, laser diode array and line detector array on the planes P_1 and P_2.

diodes, such as $(\Delta\lambda)_{stability} = 0.16$ nm. This is possible for a single laser diode at the present time.

A possible configuration of a reconfigurable optical crossbar structure is shown in Fig. 7. Multiple reflection gratings are written on a piece of z-cut LiNbO$_3$ crystal. The writing light can be derived from a dye laser or a He–Ne laser ($\lambda = 632.8$ nm), and the reading light can be derived from an array of laser diodes ($\lambda = 680.0$ nm). The wavelength difference between write-in and reading may be compensated by modifying the write-in angle according to the dispersion relation. The writing process is performed at a fixed writing angle, and multiple exposures are carried out by rotating the crystal with an angular separation equal to the required channel separation. The exposure procedure given by Strasser et al.[7] should be adopted to make the diffraction efficiencies of different channels uniform. After writing, a fixing procedure as described by Staeble et al.[8] should be used to fix the holograms in the crystal. A linear laser diode array is arranged along the nondispersion direction at the input plane. Since it is in the nondispersion direction, there would be no wavelength tunable range limitation for the separations between laser diodes, provided the total extent of the array is within the paraxial approximation. A 1-D line detector array is placed at the output plane with a separation of the line detectors equal to 2 times the channel separation, as previously given. Different laser diodes can use the same set of gratings in the crystal. In the nondispersion direction, the light beams are reflected according to the law of reflection. By independently controlling the wavelength of the laser diodes, a reconfigurable electrooptical crossbar architecture can be realized in practice.

IV. Reconfigurability with Spatial Division

Another possible architecture for reconfigurable interconnections with volume holograms is spatial division instead of wavelength tuning. A page oriented holographic setup is shown in Fig. 8. We shall apply the pinhole hologram technique as proposed by Xu et al.[9] into a photorefractive crystal. The angular selectivity of the recorded volume hologram would have a

larger storage capacity compared with thin holographic plates. The recording arrangement is shown in Fig. 8(a), in which an SLM is in the focal plane of condenser lens L_1 and the interconnection pattern masks are placed at page plane P. Object beam B is focused by L_1; after passing the SLM the beam is directed toward the recording medium. We note that the SLM is used to generate a movable pinhole that allows only the object beam to pass in one direction. In other words, the interconnection mask is illuminated by the object beam in one direction, where the pinhole of the SLM is set at a spatial location to allow the object beam to pass through. It can be seen that a set of interconnection masks can be encoded in the crystal for a given reference beam A by moving the SLM pinhole position and changing the object beam B'. Similarly, another set of interconnection pattern masks can be encoded for another reference beam A' and so on. In the readout process, a 1-D laser diode array is placed at the front focal plane Q of the collimating lens L_2 as shown in Fig. 8(b). Each diode generates a reading beam that is conjugate to a specific reference beam A. When the SLM pinhole is set at one position, a set of interconnection patterns will be diffracted and projected onto page plane P. As the pinhole position is moved, the interconnection patterns would be changed. Thus we see that the interconnections between the laser diode array and the page plane can be made reconfigurable by a programming SLM.

Unlike the wavelength tuning reconfigurable interconnection, the spatial division reconfiguration requires low wavelength selectivity (not sensitive to the wavelength). Based on the comparisons of volume holograms discussed in Sec. II, we should use transmission type holograms for interconnections. To have a higher angular selectivity, the average write-in angle (2α) should be $\sim 90°$. If the thickness of the photorefractive crystal (i.e., $LiNbO_3$) is $\sim 1-2$ cm, the bandwidth of the hologram would be much wider than the signal bandwidth as can be seen in Fig. 4(a). With reference to Fig. 3(a), the angular selectivity is found to be $(\Delta\alpha)_t = 5\lambda/d = 3.4 \times 10^{-4}$. If $3(\Delta\alpha)_t$ is used as the channel separation and assuming a 30° angular range, there would be $\sim 10^3$ useful interconnection channels in the photorefractive crystal. Due to the degeneration of Bragg diffraction, only a 1-D laser diode can be used. If the full range of $N \times N$ SLM pinholes is used, the total number of interconnection patterns would be $M = N^2 \times K$, where K is the number of channels (i.e., the number of laser diodes).

In principle, this architecture can be used for massive information storage. As an example if $N = 32$, $K = 10^3$, the total number M of patterns (or pictures) would be 10^6, which is about the capacity of a 10-h TV program. This architecture does not require a high space–bandwidth product of the SLM (only 32 × 32). The requirement of the pinhole size is about $m\lambda F/a$, where $m \times m$ is the number of pixels, a is the size of the pictures, F is the focal length of the condenser lens L_1. The size of the SLM is about $mN\lambda F/a$. For example, if $m = 512$, $N = 32$, $\lambda = 680$ nm, $F/a = 2$, the required size
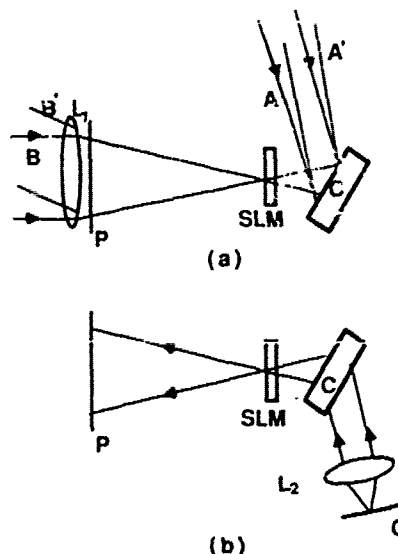

Fig. 8. Geometry for reconfigurable interconnections with spatial division: (a) recording setup; (b) reading setup. L_1, condensing lens; L_2, collimating lens; SLM, spatial light modulator; C, recording media; P, page plane; Q, laser diode plane.

of the SLM is $\sim 23 \times 23$ mm. The main advantage of this architecture is its rapid and random access to the pictures in contrast with the sequential access in optical disks.

In practice, the total number M of interconnection patterns which can be recorded in the volume hologram is basically limited by the maximum index modulation n_1 of the photorefractive material. With reference to Ref. 10, the number of interconnection patterns is about

$$M = \frac{\pi}{\eta^{1/2}} \frac{n_1}{n} \frac{d}{\lambda}. \qquad (25)$$

For example, given $\Delta n \sim 10^{-3}$, if $\eta \sim 10^{-3}$, $d = 2$ cm, and $\lambda = 680$ nm, for an ion-doped $LiNbO_3$ photorefractive crystal, the total number of interconnection patterns would be $M \sim 1.2 \times 10^3$. Recently, photorefractive fibers have been successfully grown. The fibers can be used to synthesize a large volume of photorefractive material,[11,12] which is not possible to grow in a single crystal form. The effective thickness d of the material in Eq. (25) can be expected to increase by 1–2 orders of magnitude. In addition, new mater. with higher n_1 can also be expected (e.g., organic materials).[13] Therefore, the material limitation of Eq. (25) may be alleviated in the future. On the other hand, we would like to point out that the crosstalk from third-order diffraction, scattering, internal reflections, and noise behavior of the volume holograms should be investigated before practical application of angular or wavelength multiplexing in photorefractive materials.

V. Experimental Demonstrations

The experimental setup for reconfigurable interconnections with wavelength tuning is shown in Fig. 9.
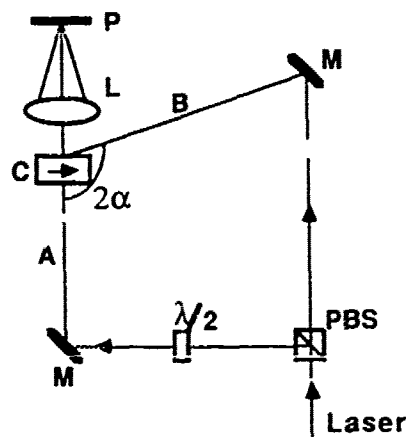
Fig. 9. Experimental setup: C, y-cut LiNbO₃; PBS, polarizing beam splitter; λ/2, halfwave plate; M, mirror; L, focusing lens; P, output plane.

Fig. 11. Reconstruction wavelength as a function of the deflected angle.

Fig. 10. Readout light spots with eight different wavelengths.

(a)

(b)

Fig. 12. Reconfigurable interconnections with spatial division: (a) reconstructed patterns on the page plane by three different reading directions with the same pinhole position; (b) reconstructed six focal spots on the SLM plane by one reading beam.

The laser beam is divided into two parts by a polarizing beam splitter. A halfwave plate was used to rotate the polarization of one beam so that the reflection hologram was formed in ordinary polarization. A piece of y-cut LiNbO₃ crystal 1 mm thick was employed as the recording medium. The write-in angle 2α of beams A and B is ~150°. Since we do not have a z-cut crystal, to have the grating vector component on the c-axis, the writing arrangement should be asymmetrical. Beam A is arranged to be perpendicular to the crystal surface. The wavelength used for writing is 632.8 nm. Eight exposures were made by rotating the crystal with a 0.8° step. A tunable dye laser is used to read the hologram in the same direction as beam B. The reconstructed light field is collected by focusing lens L onto output plane P. It is, therefore, apparent that holographic gratings can be read out one at a time by simply tuning the wavelength of the dye laser. The readout spots from eight gratings encoded in the crystal using different wavelengths are shown in Fig. 10. Figure 11 shows a linear relationship of the reconstruction wavelength vs the deflection angle. It is clear that the write-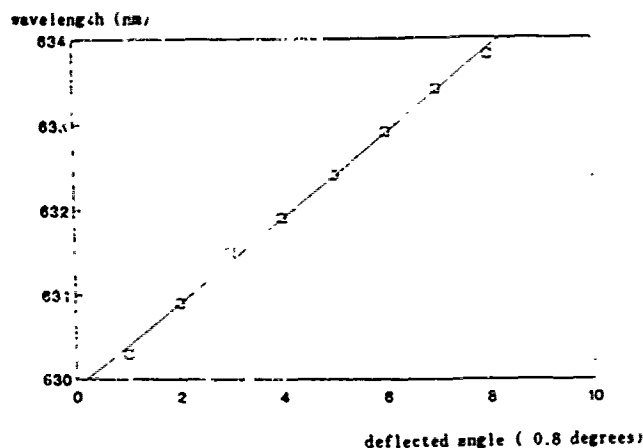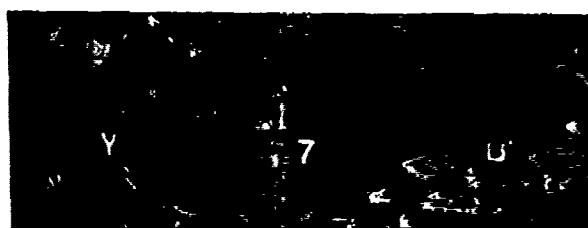in angle (i.e., 2α), calculated from the slope of the straight line using the dispersion relation of Eq. (2), is given by $2\alpha = 148.5°$.

The experimental setup for spatial division reconfigurable interconnections is basically the same as shown in Fig. 8, except that one laser beam illumination in different directions is used instead of a laser diode array. In our experiment we employed a movable pinhole as the SLM pinhole. The moving step of the pinhole is ~1 mm; the focal length of the condenser lens is 250 mm. Figure 12(a) shows three reconstructed patterns for the same pinhole position in three

different reading directions. The angular separation of the reading beams is ~1°. Figure 12(b) shows the reconstructed six focal spots at the SLM plane with a single reading direction. If these six focal spots are projected onto the page plane, six interconnection patterns would overlap. To separate the overlapping patterns, a pinhole can be used to select one focal spot at a time. Thus nonoverlapping patterns can be sequentially displayed on the page plane.

VI. Conclusion

We have analyzed the dispersion, angular, and wavelength selectivities for transmission and reflection volume holograms. The refraction on the surfaces of the crystal plays an important role in the performance of the holograms. Since the reflection holograms have higher and more uniform dispersions, it is more suitable for the purpose of reconfigurable optical interconnections with a limited tunable wavelength range. It is possible to satisfy the communication channel bandwidth if the thickness of the crystal is properly selected. In principle, a large number of reconfigurable optical interconnections with high channel bandwidth (e.g., 10^{10} Hz) can be achieved by using reflection holograms.

We have also shown that the transmission volume hologram is more suitable for the spatial division scheme. Since the transmission volume hologram is encoded in the form of pinhole holograms, the optical interconnection pattern from each laser diode illumination can easily be reconfigured by a programmable SLM. Nevertheless, for both the wavelength tuning and spatial division schemes, the maximum number of interconnections stored in a volume hologram is limited by the maximum index modulation of the photorefractive material.

References

1. J. W. Goodman, F. Leonberger, S. Y. Kung, and R. Athale, "Optical Interconnections for VLSI Systems," Proc IEEE 72, 850–866 (1984).
2. A. A. Sawchuk and B. K. Jenkins, "Dynamic Optical Interconnections for Parallel Processors," Proc. Soc. Photo-Opt. Instrum. Eng. 625, 143–149 (1986).
3. P. A. Yeh, A. E. T. Chiou, and J. Hong, "Optical Interconnection Using Photorefractive Dynamic Holograms," Appl. Opt. 27, 2093–2096 (1988).
4. F. Lin, "Optical Holographic Interconnection Network for Parallel and Distributed Processing," in Technical Digest, Topical Meeting on Optical Computing (Optical Society of America, Washington, DC, 1989).
5. E. Bradley, P. K. L. Yu, and A. R. Jonston, "System Issues Relating to Diode Requirements for VLSI Holographic Optical Interconnections," Opt. Eng. 28, 201–211 (1989).
6. H. Kogelnik, "Coupled Wave Theory for Thick Hologram Grating," Bell Syst. Tech. J 48, 2902–2947 (1969).
7. A. C. Strasser, E. S. Maniloff, K. M. Johnson, and S. D. D. Goggin, "Procedure for Recording Multiple-Exposure Holograms with Equal Diffraction Efficiency in Photorefractive Media," Opt. Lett. 14, 6–8 (1989).
8. L. Staeble, W. J. Burke, W. Phillips, and J. J. Amodei, "Multiple Storage and Erasure of Fixed Holograms in Fe-Doped LiNbO₃," Appl. Phys. Lett. 26, 182–186 (1975).
9. S. Xu, G. Mendes, S. Hart, and J. C. Dainty, "Pinhole Hologram and its Applications," Opt. Lett. 14, 107–109 (1989).
10. L. Solyman and D. J. Cooke, Volume Holography and Volume Gratings (Academic, New York, 1981).
11. J. K. Yamamoto and A. S. Bhalla, "Growth of Sr₁Ba₁₋ₓNb₂O₆ Single Crystal Fibers," Mater. Res. Bull. 24, 761–765 (1989).
12. L. Hesselink and S. Redfield, "Photorefractive Holographic Recording in Strontium Barium Niobate Fibers," Opt. Lett. 13, 877–879 (1988).
13. S. Gray, "A New Breed of Photonic Polymers," Photon. Spectra 23, 9 (1989).

APPENDIX 10.12

Binary Phase Correction

Closed-Loop Binary Phase Correction of an LCTV Using a Point Diffraction Interferometer

EDDY C. TAM, STUDENT MEMBER, IEEE, SHUDONG WU, ARIS TANONE, FRANCIS T. S. YU, FELLOW, IEEE, AND DON A. GREGORY

Abstract—A point diffraction interferometer is used to examine the phase nonuniformity of a liquid crystal television (LCTV). The resulting interference pattern is then binarized and written back onto the LCTV for phase compensation. A binary phase joint transform correlation experiment is performed using the binary phase corrected LCTV to demonstrate the applicability of this technique.

INTRODUCTION

COMMERCIALLY available liquid crystal televisions (LCTV) have been used as an inexpensive but effective spatial light modulators (SLM) for many optical signal processing applications [1]-[3]. This device can also be operated as a binary phase only SLM, by orientating the second analyzer to be perpendicular to the bisector of the two transmitted polarization states of the LCTV, as illustrated in Fig. 1 [4]. However, the nonuniform thickness of the liquid crystal display panel causes phase distortion which prohibits the direct application of the LCTV's in certain optical processing experiments that require high phase precision. Many researchers have immersed the display panel in a liquid gate filled with index matching oil, thus removing the external phase distortion. Casasent *et al.* [5] and Yu *et al.* [6] have also proposed the use of a phase conjugation method for alleviating this problem.

An article by Kim *et al.* [7] suggested correcting the phase distortion of an SLM by writing the binary phase distortion function back onto the device. In this letter, a closed-loop binary phase correction technique is demonstrated which uses a point diffraction interferometric technique [8] instead of a Mach–Zehnder interferometer as mentioned in [7]. A closed-loop and on-line architecture is introduced so the optimum correction function can be generated.

POINT DIFFRACTION INTERFEROMETER

The point diffraction interferometer was first proposed by Smart and Strong [8] as a technique for measuring the phase variation across a wavefront, the principle of which is illustrated in Fig. 2. The wavefront to be examined is brought to a focal spot by a converging lens, and is then transmitted
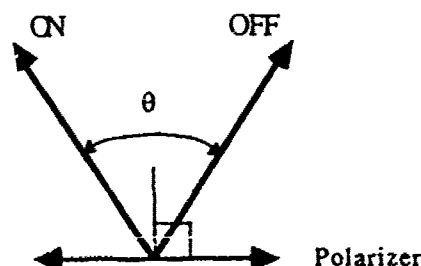
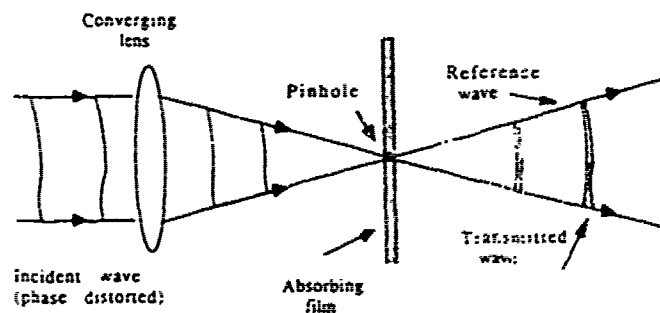Fig. 1. Orientation of the polarizer to achieve bipolar phase only modulation.

Fig. 2. Principle operation of a point diffraction interferometer.

with reduced amplitude through an absorbing film placed at the focal plane. The film has a small transparent pinhole acting as a point source which produces a spherical reference wavefront. The wavefront under examination then interferes with the reference wave when being recorded by a detector placed behind the absorbing film. The usual focus shift and tilt of the reference wavefront are made possible by displacing the pinhole longitudinally and laterally, respectively, from the focal spot of the optical system. However, in the application discussed here, it is desirable to place the pinhole at the center of the focal plane since a tilt or focal shift would introduce additional interference fringes that cannot be completely compensated for by a binary phase correction function. Since the point diffraction interferometer is a common-path interferometer the fringes are stable against vibration, and wide-band light sources can be used. The simplicity of the point diffraction interferometer architecture is also one of its major merits.

BINARY PHASE CORRECTION

Let a unity amplitude function $h(x, y)$ represent the nonuniform phase distribution on the LCTV's display panel. If the complex conjugate of this function, $h^*(x, y)$, is displayed on the LCTV, then the overall transmittance function would be
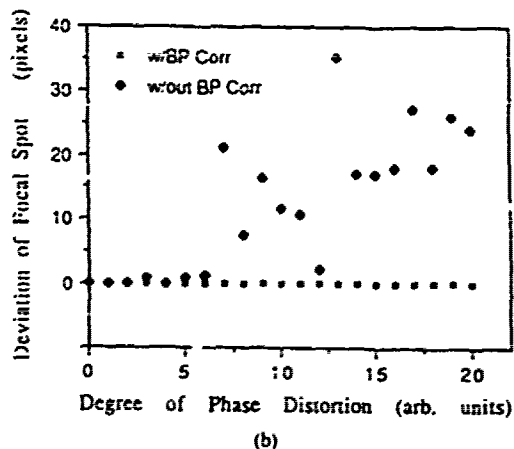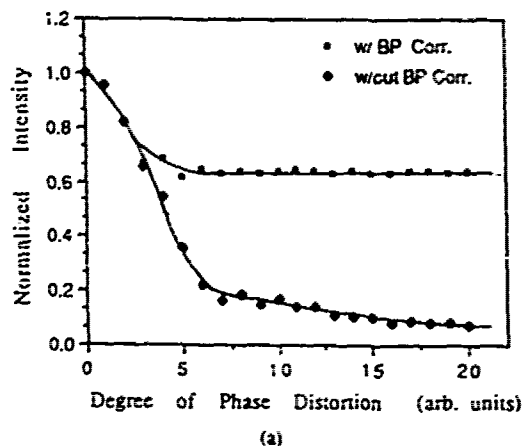
(a)

(b)

Fig. 3. (a) Computer-simulated result for binary phase correction of a phase distorted surface. Each surface is made from a circular surface (radius = 40 pixels) enclosed in a 256 × 256 image array. A set of Gaussian distributed phase angle values (standard deviation = σ in π rad. unit) are first generated as seeds. The seed angles are then assigned to every 5 × 5 array on the circular surface. Finally, a neighborhood averaging filter is applied to generate the distorted surface. Degree of phase distortion as labeled on the x-axis equals to 0.1 σ. The y-axis labels the normalized peak intensity of the plane wave response. When the degree of phase distortion is below 5 units, the phase distortion on the surface basically has a standard deviation less than $\pi/2$: therefore, the binary phase correction scheme has little effect in restoring the surface. (b) Deviation of the peak from the origin as a function of the degree of phase distortion. Angle of deviation of the peak can be related to the deviation distance N (in pixel units) by the equation $\delta\theta = (N/256)(\lambda/\delta x)$, where δx is the width of a pixel.

$h(x, y)h^*(x, y) = 1$, i.e., phase distortion of the display panel would be removed. However, the LCTV, operating in a phase modulation mode, can only generate two phase values (0 and π phase). Therefore, the proper phase correction function to be used is the binarized function given by

$$h_{bi}^*(x, y) = \begin{cases} +1 & \text{if Re } \{h^*(x, y)\} > 0: \\ -1 & \text{otherwise.} \end{cases} \quad (1)$$

Conceptually, this is equivalent to adding a zero phase to the region of the SLM if the phase distortion is in the first and fourth quadrant, and a π phase to the region if the phase is in the second and third quadrant. Although perfect phase compensation cannot be achieved, phase distortion of the display panel after correction is confined to the region from $-\pi/2$ to $-\pi/2$ such that all pixel elements of the LCTV make constructive contributions to the output.
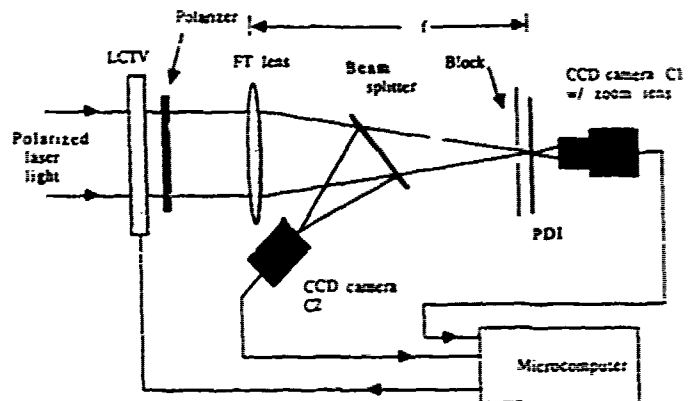

Fig. 4. Experimental setup for a closed-loop binary phase correction system using a point diffraction interferometer. The polarizer is oriented at the proper angle to achieve binary phase only modulation.

To investigate the effectiveness of this binary phase correction technique, a computer simulation was carried out. A set of randomly phase-distorted surfaces characterized by different levels of phase distortion were first generated, and their plane wave responses were computed by using an FFT routine. The intensity and position of the focal peak at the Fourier plane were recorded as a function of the degree of phase distortion. The lower curve in Fig. 3(a) shows that the peak intensity decreases rapidly as the degree of phase distortion increases. Likewise, the location of the peaks also shift randomly as shown in Fig. 3(b). The binary phase correction scheme is then employed, and the restored peak intensity is plotted in the upper curve of Fig. 3(a). It shows that the restored peak intensity is improved up to 65% of its original value even when the input function is under severe phase corruption. It can also be seen in Fig. 3(b) that the binary phase correction scheme successfully restores the peak back to its proper position.

EXPERIMENTAL DEMONSTRATIONS

Fig. 4 shows the experimental setup for implementing this proposed technique. A zoom lens is placed behind the absorbing film so that the LCTV panel can be mapped accurately onto the CCD camera C1. When the LCTV is turned off, the interference pattern recorded by C1 is given by

$$I(x, y) = |h(x, y) + C|^2 = C_1 + C_2 \text{ Re } \{h(x, y)\} \quad 2)$$

where C represents the amplitude of the reference wave due to the pinhole source on the absorbing film. Since Re $\{h(x, y)\}$ and Re $\{h^*(x, y)\}$ are equivalent, the binary phase correction function $h_{bi}^*(x, y)$ can be obtained from $I(x, y)$ with a proper threshold. To observe the phase correction result when $I_{bi}(x, y)$ is displayed on the LCTV, a beamsplitter is introduced so that the focal spot peak intensity can be recorded on CCD camera C2. Thus, the performance of the phase correction function can be evaluated as a function of the focal spot intensity. By adjusting the threshold level of the binarization process using this on-line and closed-loop architecture, the optimum binary phase correction function can be obtained, i.e., $I_{bi}(x, y) - h_{bi}^*(x, y)$.

In the experiment, an Epson color LCTV model ET708 was used. The dimensions of the display panel are 50 mm × 65
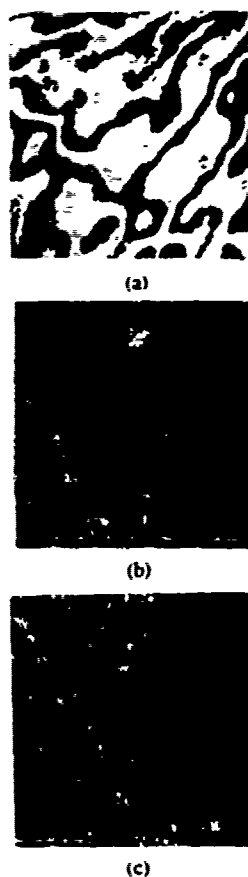
(a)

(b)

(c)

Fig. 5. (a) Correction function employed for the LCTV under test. (b) Plane wave response of the phase distorted LCTV before binary phase correction, and (c) after correction.

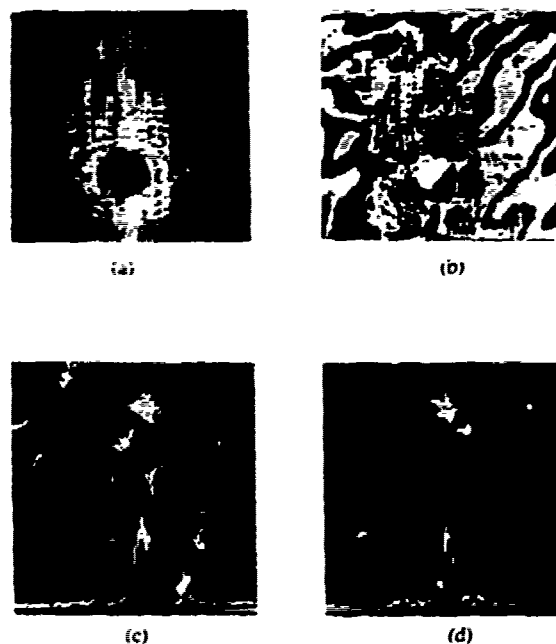(a)   (b)

(c)   (d)

Fig. 6. (a) A binarized joint transform power spectrum obtained at the Fourier plane of a JTC whose input is two identical images. (b) The Exclusive OR result between the power spectrum in (a) and the binary correction function in Fig. 5(a). (c) Correlation output when (a) is displayed on the LCTV. The dc is at the center of the photograph, while the two peaks on the sides are the correlation functions. (d) Correlation output when (b) is displayed on the LCTV

mm. the resolution is 238 pixels × 442 pixels. The modulation angle between the two polarization states. given as θ in Fig. 1. was measured to be about 46° in this experiment setting. The aperture used in the point diffraction interferometer was made by burning a pinhole ($\approx 5 \mu m$) on a $\approx 2\%$ transmittance film. Fig. 5(a) shows the optimum binary phase correction function thus generated. The focal spots recorded by C2 before and after phase correction are shown in Figs. 5(b) and (c). respectively. It can be seen that the phase compensation produces a sharper plane wave response whose intensity is almost one-third higher than that without phase compensation.

A binary image $f(x, y)$ can also be written onto the binary phase corrected LCTV for image processing. In this case. the final phase image displayed on the LCTV is given by $f(x, y)h_c^*(x, y)$. Since both of these functions are binary in nature. the multiplicative operation can be replaced by an Exclusive OR operation which facilitates digital implementation. To demonstrate that the binary phase corrected LCTV can be used in binary-input image processing applications, a binary phase only joint transform correlation [9] experiment was performed. An LCTV (immersed in a liquid gate) was used at the input plane to display two identical nonbinary amplitude modulated input images. and a CCD camera was placed at the Fourier plane to record the joint transform power spectrum. The power spectrum thus obtained was then binarized by using a microcomputer and displayed on the binary phase corrected LCTV for the second Fourier transformation. The

binarized joint transform power spectrum is shown in Fig. 6(a). while Fig. 6(b) shows the Exclusive OR result between the spectrum and the phase correction function. Fig. 6(c) and (d) shows the output correlation results before and after the binary phase correction function is applied. Results are similar to those shown in Fig. 5: sharper correlation peaks whose intensities are about one-third higher than those without binary phase correction are observed.

The same phase correction principle can be applied to correct the phase distortion of an optical signal processing system instead of the LCTV itself. The TV panel is used as a real-time addressable binary phase filter solely for phase correction purposes. To illustrate this idea. the LCTV display was immersed into a home-made liquid gate. The nonuniform phase distortion of the LCTV was corrected by the index matching oil. However. the undesirable curvature of the liquid gate's optical window introduced an astigmatic aberration. whose interference pattern. obtained using the point diffraction interferometer. is shown in Fig. 7(a). The focal spot structures. recorded at the proper focal plane before and after binary phase correction by the LCTV. are shown in Fig. 7(b) and (c). respectively. The programmable phase filter. i.e.. the LCTV. effectively removed the aberration. and a sharp focal spot was observed at the correct focal plane.

CONCLUSIONS

Binary phase correction of an LCTV using a point diffraction interferometer has been demonstrated. The use of a point diffraction interferometer offers the simplicity and advantages of a common-path interferometer for examining the phase nonuniformity of the LCTV. while the on-line and closed-
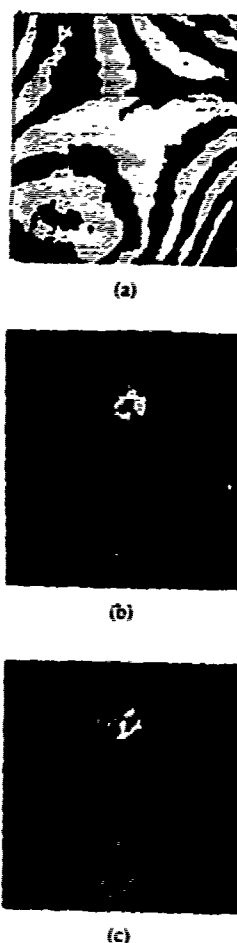
Fig. 7. (a) Binary correction function for an astigmatic aberrated optical system. (b) Plane wave response of the aberrated system before binary phase correction. and (c) after correction.

loop architecture is capable of generating the optimum phase correction function. Both computer simulations and experimental results demonstrate the effectiveness of the binary phase correction scheme. The results obtained from the bipolar phase only joint transform correlation reveal the feasibility of this phase correction technique in actual applications. Mention must be made that this binary phase correction scheme can be used to compensate for the phase distortion of a general optical signal processing system. Other phase modulating real-time addressable SLM's which have a higher contrast ratio and a larger space-bandwidth product. such as the magnetooptic SLM's [10] or liquid crystal light valves [11]. may be employed to produce even better results.

REFERENCES

[1] H. K Liu. J. A. Davis. and R. A. Lilly. "Optical date processing properties of a liquid crystal television spatial light modulator." Opt Lett.. vol. 10. pp. 635-637. 1985.

[2] D. A. Gregory. "Real-time pattern recognition using a modified liquid crystal television in a coherent optical correlator. " Appl. Opt . vol. 25. pp 467-469. 1986.

[3] F. T. S. Yu. S. Jutamulia. and X. L. Huang. "Experimental application of low-cost liquid crystal TV to white-light optical signal processing." Appl. Opt . vol. 25. pp. 3324-3326. 1986.

[4] J. A. Davis. R. A. Lilly. K. D. Krenz. and H. K. Liu. "Applicability of the liquid crystal television for optical data processing." Nonlinear Optics and Application. Proc. SPIE, vol. 613. pp. 245-254. 1986.

[5] D. Casasent and S. F. Xia. "Phase correction of light modulators." Opt. Lett.. vol. 11. pp. 398-340. 1986.

[6] F. T. S. Yu. S. Wu. and A. W. Mayers. "Applications of phase conjugate to a joint transform correlator." Opt. Commun.. vol. 71. pp. 156-160. 1989.

[7] H. M. Kim. J. W. Jeong. M. H. Kang. and S I. Jeong. "Phase correction of a spatial light modulator displaying a binary phase-only filter." Appl. Opt.. vol. 27. pp. 4167-4168. 1988.

[8] R. N Smart and J. Strong. "Point diffraction interferometer. " J Opt. Soc. Amer.. vol. 62. p. 737. 1972.

[9] F. T. S. Yu and T. Nagata. "Binary phase only joint transform correlator." Micro. Opt. Tech. Lett.. vol. 2. pp. 15-19. 1989.

[10] D. Psaltis. E. G. Paek. and S. S Venkatesh. "Optical image correlation with a binary spatial light modulator." Opt. Eng . vol. 23. pp. 698-704. 1984.

[11] J. A. Davis. G. M. Heissenberger. R A. Lilly. D. M. Cottrell. and M. F Brownell. "High efficiency optical reconstruction of binary phase-only filters using the Hughes liquid crystal light valve." Appl Opt . vol. 26. pp. 929-933. 1987.

APPENDIX 10.13

Image Classification

IMAGE CLASSIFICATION BY THE KITTLER-YOUNG TRANSFORM USING A JOINT TRANSFORM CORRELATOR

John X. Li and Francis T. S. Yu
Department of Electrical Engineering
The Pennsylvania State University
University Park, Pennsylvania 16802

Don A. Gregory
U S. Army Missile Command
Redstone Arsenal, Alabama 35898

KEY TERMS

ABSTRACT

This article will discuss a two-category image classification which has been performed on a joint transform correlator (JTC) by using the Kittler–Young transform. The property of optimally utilizing the class means and variances in the Kittler–Young transform has enabled a high space-bandwidth-product operation. The digitizing effect of the liquid crystal television has been utilized to increase the correlation peaks.

The application of statistical pattern recognition introduces effective feature extraction to optical systems [1]. In statistical pattern recognition, principal component analysis [2] is widely used (e.g., Clafic [3], Selfic [2], the Fukunaga–Koontz (FK) transform [4], and the Kittler–Young (KY) transform [5]). The FK transform has been implemented by Leger and Lee for 64 × 64 images on the Vander Lugt correlator [6]. The first three algorithms mentioned above have a common characteristic. The feature selection criteria are based on the corresponding eigenvalues generated from the principal component analyses. However, the KY transform goes one step further in using the eigenvalues, such that a larger space-bandwidth-product (SBP) process can be performed in pattern recognition. This article presents the implementation of the KY transform on the joint transform correlator (JTC) [7, 8] to perform a two-category image classification. The JTC implementation of the KY transform can eliminate the pre-process of matched filter synthesis, which is required in the Vander Lugt correlator. Furthermore, the resolution requirement at the frequency plane in JTC is much lower than the Vander Lugt correlators, such that the direct application of spatial light modulators (SLM) at the frequency plane is possible. As a result, high SBP image classification is achieved.

The KY transform is a linear transformation to extract features from images. The transform matrix for the KY transform is constructed in a training process. Training images of $n \times n$ pixels used in the training process are lexicographically scanned and stored as column vectors $x^{(i)}j$ of dimension N ($= n^2$), where indices i ($= 1, 2, \ldots, K$) and j ($1, 2, \ldots, M$) are the class and image indices, respectively. The first step in constructing the KY transform matrix is to diagonalize the covariance matrix C [5] of the training images, in which the eigenvector matrix Φ, and the eigenvalue matrix Λ of the covariance matrix C are obtained. Each eigenvalue λ_k is the average class variance of the corresponding feature. In order to normalize the class variances of the features, a whitening transform is performed [9] in the second step. After two successive linear transforms with transform matrices Φ and $\Lambda^{-1/2}$, image vectors x are mapped onto feature vectors v. The covariance matrix in y space becomes the identity matrix with all the average class variances normalized to unity. In the final step, the correlation matrix R, [5] in y space is diagonalized, and the overall KY transform matrix T is thus formed:

$$T = \Phi\Lambda^{-1/2}B = (t_1, t_2, \ldots, t_{KM}), \qquad (1)$$

where B is the eigenvector matrix of the correlation matrix R, and t_k is the kth column vector in the matrix T. Image vectors x can be mapped into the final feature space f:

$$f = Tx. \qquad (2)$$

The KY transform generates features which are more discriminative. The eigenvalues, generated from the principal component analysis of the correlation matrix, are usually composed of two parts that are proportional to class means and class variances, respectively. The effects of these two parts on the feature's discriminative power are quite opposite [5]. The conflict of the class means and variances in the eigenvalues, which is common in the other algorithms, is overcome in the KY transform as follows. The contribution of the class variance in f space to the new eigenvalues r_k ($k = 1, 2, 3, \ldots$, KM) of the second principal component analysis remains the same as those in the y space. Therfore, the eigenvalues r_k are solely related to the class means:

$$r_k = 1 + \sum_{i=1}^{k} P_i\chi_{ki}^2, \qquad k = 1, 2, 3, \ldots, KM. \qquad (3)$$

χ_{ki} is the kth component of the ith class mean in f space. On the other hand, the feature selection criterion δ_k ($k = 1, 2, 3, \ldots, KM$) is defined as the weighted summation of the average mutual distance between the class means:

$$\delta_k = \sum_{i=1}^{k} P_i \sum_{l=1}^{K} P_l(\chi_{ki} - \chi_{kl})^2, \qquad k = 1, 2, 3, \ldots, KM. \qquad (4)$$

and can be reduced [5] to

$$\delta_k \approx r_k - 1. \qquad (5)$$

As a result, the feature selection criterion can be based on the eigenvalues r_k, which are solely related to the class means. Therefore, with the eigenvalues r_k arranged in descending
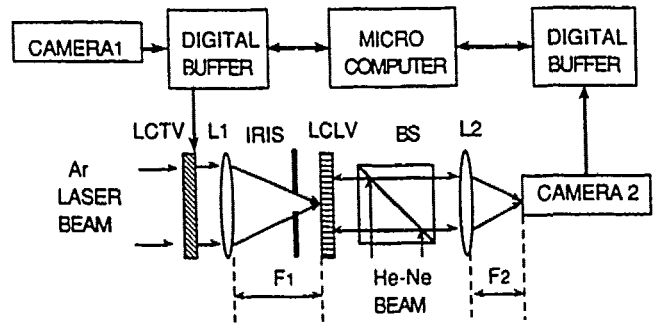

Figure 1   Joint transform correlator consisting of LCTV liquid crystal television, L1 transform lens, LCLV liquid crystal light valve, BS beamsplitter, and L2 transform lens

order, image classification which uses the first m features (where $m < KM$) is likely to be more accurate than the other algorithms. Moreover, a higher SBP process in image classification is possible simply because the classification error rate is proportional to the size of images N [10].

The implementation of the KY transform is conducted on a programmable JTC, shown in Figure 1, for a two-category image classification. In the training process, 10 128×128 training images ($M = 10$) of each class are used. In order to limit the bias between the training-image error rate and the test-image error rate [11], the first two features of the KY transform, f_1 and f_2, are selected for the classification ($m = 2$). In the two-dimensional implementation on the JTC, the feature extraction can be represented by the following:

$$f_k = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} t_{k\alpha\beta} x_{\alpha\beta}, \qquad k = 1, 2, \ldots, m, \qquad (6)$$

where $t_{k\alpha\beta}$ is the $n \times n$ array format of the column vector t_k in Eq. (2) and $x_{\alpha\beta}$ is the input image. The three-dimensional KY transform matrix T can be partitioned into m submatrices $t_{k\alpha\beta}$, which are called the KY basis images. The basis images 1 and 2 are shown in Figure 2(a) and (b), respectively. The KY features f_k of a target image are obtained at the correlation center in the output plane of the JTC, while the target and the kth KY basis image are displayed, side by side, in the input plane. A 3-in liquid crystal television (LCTV) of 220×240 pixels is used as the input device, which is submerged in a liquid gate to make the phase distribution across the aperture uniform. A 48-in lens has to be used as the front Fourier transform lens in order to write the essential grating in the joint power spectrum on the liquid crystal light valve (LCLV), which has a cutoff frequency of 40 lines/mm. The shift invariant property of the JTC cannot be fully utilized, due to the limited number of displaying elements in the LCTV. The LCTV also affects the contrast ratio of the JTC. However, by reducing the frequency of the ac power supply for the LCLV to 1 kHz, the dc spot in the joint power spectrum

(a)   (b)

Figure 2   (a) The KY basis image 1, and (b) basis image 2

diminishes so that the light leakage through the space between the electrodes is blocked at the frequency plane. Therefore, a contrast ratio of 30 : 1 is achieved in the JTC. The addressing and relaxation times of the JTC are 80 and 70 ms, respectively, of which the LCLV contributes a major part. A television frame rate (30 frames/sec) operation can be achieved by replacing the CdS LCLV with the Si LCLV [12].

Increased cross-correlations are obtained in the output plane by taking advantage of the digitizing effect of the LCTV. Due to the digitization of the input image by the LCTV, the joint power spectrum takes a lattice form, where each dot in the lattice carries the same information content. The cross correlation between a target and a KY basis image in the output plane is given by

$$o(x, y) = \{[t(x, y) \otimes b(x, y)] * \delta(x - 2\Delta x, y)\}$$

$$\times \left\{ \sum_{m, n} \exp[-i2\pi(m'x/s_t + n'y/s_t)] \right\}, \quad (7)$$

where the symbols $\otimes$ and $*$ denote the two-dimensional correlation and convolution operation, respectively; $t(x, y)$ and $b(x, y)$ represent the target and the KY basis image, respectively; s_t and s_t are the sampling periods of the LCTV in the x and y directions; $2\Delta x$ is the separation between the two images; λ is the wavelength of the readout coherent light; and f is the focal length of the lens L2. By properly adjusting the iris in front of the LCLV, the magnitude of the extra phase term in Eq. (7) may exceed unity such that substantially increased correlation peaks can be obtained. An example of the correlation output is shown in Figure 3, where zero-order diffraction (i.e., dc) is partially blocked. Moreover, a quasi-

monochromatic and partial coherent can be employed in readout to reduce the coherent noise.

The two KY features, f_1 and f_2, are extracted from a target image in two separate joint transform correlation operations. The same process is repeated 32 times to generate the two KY features for all 32 images (20 of which are sample images, and 12 of which are test images). The KY features of all the images are plotted in Figure 4. A solution linear classifier is obtained by using the gradient descent method [13] and is also plotted in Figure 4. The results show that there is no error in the two-category classification of the 32 128 × 128 images. Also, the image classification by the KY transform can be easily extended to multicategory cases.

REFERENCES

1. Q. Tian, Y. Fainman, and Sing H. Lee, *J. Opt. Soc. Am.*, Vol. 5, 1988, pp. 1670–1682.
2. S. Watanobe, *Proc. 4th Prague Conf. on Information Theory*, 1965.
3. S. Watanobe et al., *Computers and Information Sciences*, Academic, New York, 1967, Vol. 2.
4. K. Fukunaga and W. L. G. Koontz, *IEEE Trans. Comput.*, Vol. C-19, 1970, pp. 311–318.
5. Josef Kittler and Peter C. Young, *Pattern Recognition*, Vol. 5, 1973, pp. 335–352.
6. J. R. Leger and S. H. Lee, *J. Opt. Soc. Am.*, Vol. 72, 1982, pp. 556–564.
7. F. T. S. Yu and J. E. Ludman, "Joint Fourier Transform Processor," *Microwave Opt. Technol. Lett.*, Vol. 1, 1988, p. 374.
8. F. T. S. Yu and X. J. Lu, "A Real-Time Programmable Joint Transform Correlator," *Opt. Commun.*, Vol. 52, 1984, p. 10.
9. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic, New York, 1972
10. Q. Tian, M. Barbero, Zu-han Gu, and Sing H. Lee, *Opt. Eng.*, Vol. 25, 1986, pp. 834–840.
11. Donald H. Foley, *IEEE Trans. Inf. Theory*, Vol. IT-18, 1972, pp. 618–626.
12. Keyvan Sayyah, Murray S. Welkowsky, Philip G. Reif, and Norman W. Goodwin, *Appl. Opt.*, Vol. 28, 1989, pp. 4748–4756.
13. Richard O. Duda and Peter E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

Figure 3 Optical correlation output showing two correlation spots (marked with arrows) and the partially blocked dc spot in the middle

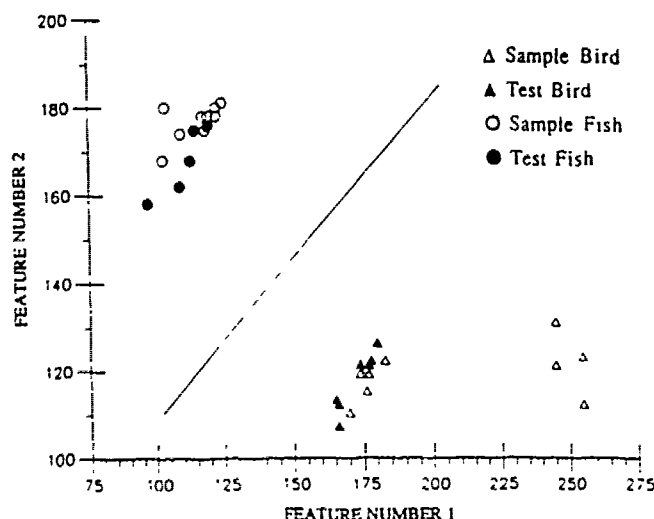**Figure 4** Image classification in the Kittler–Young feature space

APPENDIX 10.14

IPA Neural Network

Neural network model using interpattern association

Taiwei Lu, Xin Xu, Sudong Wu, and Francis T. S. Yu

This paper investigates a neural network model—interpattern association (IPA) model—in which the basic logical operations are used to determine the interpattern association (i.e., association between the reference patterns), and simple logical rules are applied to construct tristate interconnections in the network. Computer simulations for the reconstruction of similar English letters embedded in the random noise by the IPA model have shown improved performance compared with the Hopfield model. A 2-D hybrid optical neural network is used to demonstrate the usefulness of the IPA model. Since there are only three gray levels used in the interconnection weight matrix for the IPA model, the dynamic range imposed on a spatial light modulator is rather relaxed, and the interconnections are much simpler than the Hopfield model.

I. Introduction

Neural networks have shown effectiveness in pattern recognition.[1-] For recognition of a given set of reference patterns, there are two schemes to construct the associative memory matrix (i.e., interconnection weight matrix). The first is the so-called intrapattern association, which emphasizes the association of elements within each reference pattern. For example, in the Hopfield model, the outer products of the reference patterns are added to form the interconnection weight matrix. This type of approach may create an unstable or ill-conditioned network if the reference patterns are not independent of each other. The second scheme is an interpattern association (IPA) for which the interconnection weight matrix is constructed by emphasizing the association between reference patterns. Where the reference patterns are similar to each other (e.g., human faces, fingerprints, handwritten characters), the special features of each pattern become very important in pattern recognition. Therefore, it is necessary to consider the relationships between the special and common features among the reference patterns in constructing the interconnection weight matrix.

We present a neural network model based on the association between reference patterns. This model first determines the common and special features

among the reference patterns by applying a logical operation. Two equivalent logical rules are presented in Sec. II to construct the excitatory and inhibitory interconnections in the network. In Sec. III, the IPA model is compared with the Hopfield model. Computer simulations for the reconstruction of similar English letters in the random noise have shown that the IPA model performs better than the Hopfield model. In Sec. IV, a 2-D hybrid optical neural network is used to implement this new model. The experimental results show that the IPA model indeed performs more effectively than the Hopfield model.

II. Interpattern Association Model

The IPA is important in the construction of the associative memory matrix. The information obtained from special areas (i.e., areas occupied by one pattern) is important than from common or overlapped areas.

Figure 1 shows an example of a reference set that consists of three overlapping patterns A, B, and C in the pattern space. These patterns can be divided into seven subspaces. I, II, and III are the special areas of patterns A, B, and C, respectively. IV, V, and VI are the common areas of A and B, B and C, C and A, respectively. VII is the common area of A, B, and C. The rest can be defined as an empty space Φ. These areas can be expressed by the following logic functions:

$$\text{I} = A \wedge (\overline{B \vee C}),$$
$$\text{II} = B \wedge (\overline{A \vee C}), \quad \text{V} = (B \wedge C) \wedge \overline{A},$$
$$\text{III} = C \wedge (\overline{A \vee B}), \quad \text{VI} = (C \wedge A) \wedge \overline{B},$$
$$\text{IV} = (A \wedge B) \wedge \overline{C}, \quad \text{VII} = (A \wedge B \wedge C) \wedge \overline{\Phi}, \quad (1)$$

where $\wedge$, $\vee$, and $\overline{}$ stand for the logic AND, OR, and NOT operations, respectively.

The authors are with Pennsylvania State University, Department of Electrical Engineering, University Park, Pennsylvania 16802.
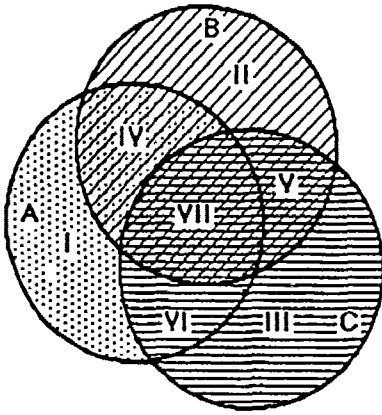
Fig. 1. Common and special areas of three reference patterns.
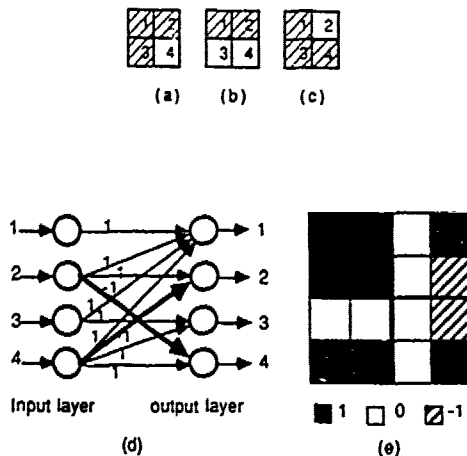

Fig. 2. Example of constructing an IPA neural network: Three 2 ×
2 reference patterns: A (a), B (b), C (c), tristate neural network (d),
interconnection weight matrix (e).

With the above information, we can start building the interconnections for a single layer neural network. To illustrate the concept of IPA in this example, each pixel in the reference pattern space is mapped to each input/output neuron.

When an input neuron in area VII is on (i.e., the input of this neuron is 1), it implies that there is an input but not whether it belongs to any pattern. Thus this neuron can only excite the output neurons within area VII and have no connection with the output neurons in other areas.

When an input neuron in area V is on, it implies that the input pattern is not A but not whether it belongs to pattern B or C. Thus this neuron must excite the output neurons in areas V and VII but inhibit the output neurons in area I, similarly for input neurons in area IV or VI.

For the case when an input neuron is on in area I, it implies that pattern A appears at the input end. Therefore, this neuron can excite all the output neurons in pattern A (i.e., areas I, IV, VI, and VII) but inhibit the output neurons in areas (B ∨ C) ∧ $\bar{A}$ (i.e., areas II, III, and V). Similarly, for the input neurons in area II or III, they must excite output neurons in

area B or C and inhibit neurons in areas I, III, and VI or areas I, II, and IV, respectively.

In view of Eq. (1), the descriptive logic can be summarized in the following:

Rule A:

Assume that a subarea X is represented by the following expression:

$$X = P \wedge \bar{Q}, \tag{2}$$

where

$$P = p_1 \wedge p_2 \wedge \ldots \wedge p_n, \tag{3}$$

$$Q = q_1 \vee q_2 \vee \ldots \vee q_m. \tag{4}$$

Here $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_m$ are reference patterns, n and m are two positive integers, and $n + m = M$ is the total number of reference patterns. The input neurons in area X must excite (i.e., having positive connections with) all the output neurons in area P, inhibit (i.e., having negative connections with) all output neurons in area Q ∧ $\bar{P'}$ where P' is defined by

$$P' = p_1 \vee p_2 \vee \ldots \vee p_n, \tag{5}$$

and have no connection with the output neurons in the remaining areas. For simplicity, the connection strengths (i.e., weights) are defined as 1 for positive connections, −1 for negative connections, and 0 for no connection. Thus the IPA neural network can be constructed in a simple three-state structure.

To illustrate further the construction of the IPA model, A, B, and C are assumed as three 2 × 2 array patterns shown in Figs. 2(a), (b), and (c), respectively. The pixel–pattern relationship is given in Table I. It is apparent that pixel 1 represents the common feature of A, B, and C, pixel 2 is the common feature of A and B, pixel 3 is the common feature of A and C, and pixel 4 represents the special feature of C.

By applying Rule A to the three reference patterns A, B, and C, a tristate neural network can be constructed, as illustrated in Fig. 2(d). This is a one-layer neural network with four input neurons and four output neurons. Each neuron is matched to one pixel of the reference patterns. For example, the first input neurons corresponds to pixel 1 of the input pattern and excites only the first output neuron. The second input neuron (the corresponding pixel belongs to patterns A and B) excites both the first and second output neu-

Table I. Pixel-Pattern Relationship of Three Reference Patterns

Pixel / Pattern	1	2	3	4
A	1	1	1	0
B	1	1	0	0
C	1	0	1	1

rons, while inhibiting the fourth output neuron, which belongs to the special area of pattern C.

It is interesting to analyze the structure of the IWM. Since the reference patterns are in 2-D form, the weight matrix becomes a 4-D matrix. We can partition the 4-D IWM into a 2-D submatrix array,[5] as illustrated in Fig. 2(e). The IWM can be divided into four blocks. Each block corresponds to one output neuron. The four elements in one block represent the four neurons in the input end. For example, since all four elements in the upper left block have a value of 1, it indicates that either one of the four input neurons can excite the first output neuron. In the upper right block, as another example, the first and third elements are 0, the second element has a value of 1, and the fourth element is −1. From this example we can determine that the first and third input neurons have no connection to the second output neuron, the second input neuron excites the second output neuron, and the fourth input neuron inhibits the second output neuron.

To simplify the algebraic operations, an equivalent rule, Rule B, is developed as follows. It can be used to construct the IWM by examining the pixel–pattern relationships.

Rule B:

Let us define $D_{l,i}$ as a 2-D matrix that corresponds to the 2-D array in Table I, where l and i denote the row and column numbers. Let d_i be the number of patterns that are bright (i.e., in state 1) at the ith pixel. then it can be determined by summing the elements in the ith column of Table I, i.e.,

$$d_i = \sum_{l=1}^{M} D_{l,i}. \tag{6}$$

Let us also define k_{ij}:

$$k_{ij} = \sum_{l=1}^{M} D_{l,i} D_{l,j}, \tag{7}$$

which is the sum of the product of columns i and j in Table I. Then we can construct an IPA neural network by applying the following logical rules:

(1) If $k_{ij} = \min(d_i, d_j)$,

when $d_i < d_j$, pixel i must excite pixel j, but pixel j must not excite pixel i;

when $d_i = d_j$, pixels i and j must excite each other;

when $d_i > d_j$, pixel j must excite pixel i, but pixel i must not excite pixel j;

(2) If $0 < k_{ij} < \min(d_i, d_j)$, pixels i and j have no connection to each other.

(3) If $k_{ij} = 0$,

when $d_i \neq 0$, and $d_j \neq 0$, pixels i and j must inhibit each other;

when $d_i = 0$ and/or $d_j = 0$, pixels i and j must have no connection to each other.

We stress that these rules for IWMs are simple and straightforward, which is suitable for computer calculations. In fact, the IWMs computed by Rule B are equivalent to those obtained by Rule A.

III. Comparison with the Hopfield Model

It is the differences rather than the similarities among patterns that are used for pattern recognition. For example, the outline of the eyes, nose, and mouth are common features in all human faces. People distinguish different persons by the differences rather than the detail in the faces.

Similar to some other neural network algorithms, the Hopfield model constructs the IWM by correlating the elements within each pattern, however, ignoring the relationships among the reference patterns. The IWM T of the Hopfield model for three reference patterns A, B, and C can generally be expressed as

$$T = AA^T + BB^T + CC^T - 3I, \tag{8}$$

where T is the transpose of the vectors and I is the unit matrix, which makes the weight matrix zero diagonal.

If input pattern A is applied to the neural system, the output would be

$$V = TA$$

$$= A(A^T A) + B(B^T A) + C(C^T A) - 3A, \tag{9}$$

where $A^T A$ represents the autocorrelation of pattern A, while $B^T A$ and $C^T A$ are the respective cross-correlation between A and B, and A and C. If the differences between A, B and C are sufficiently large, the autocorrelation of A, B or C would be much larger than the cross-correlation between them, i.e.,

$$A^T A \gg B^T A, \qquad A^T A \gg C^T A. \tag{10}$$

Notice that from Eq. (9) pattern A has a larger weighting factor than patterns B and C, for which patterns B and C can be considered noise. By choosing the proper threshold value, pattern A will be constructed at the output end of the neural network.

On the other hand, if A, B, and C are very similar, Eq. (10) no longer holds; since the weights of patterns B and C are comparable with those of A, patterns B and C can no longer be considered noise. Thus the threshold value for the Hopfield model cannot be easily defined; the Hopfield model thus becomes unstable.

Computer simulations of a 2-D neural network with 8×8 array neurons at the input and output ends have been conducted for both Hopfield and IPA models. The reference patterns considered are the twenty-six capital English letters in a sequence based on the similarities of the , B,P,R,F, ..., and each letter occupies an 8×8 array pixel.

Figure 3 shows the error rates as a function of reference patterns for Hopfield and IPA models with noisy inputs. The input SNR for computer simulation is ~ 7 dB. We notice that the Hopfield model becomes unstable when patterns B, P, and R are stored in the IWM, whereas the IPA model is quite stable with twelve stored letters. Comparison is also made for noiseless input. In this condition, the IPA model can produce correct results for all twenty-six stored letters in the IWM, whereas the Hopfield model starts making significant errors when the number reference patterns are increased to 3.
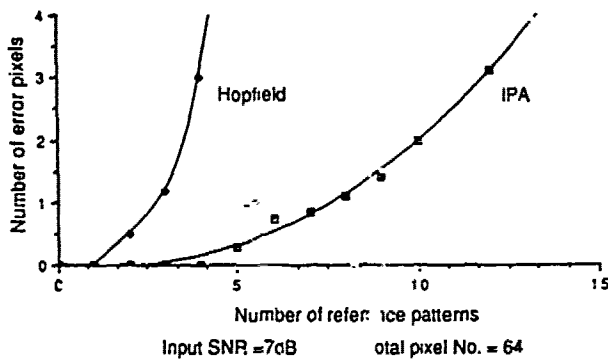
Fig. 3. Comparison of the IPA and Hopfield models.

Input SNR = 7 dB   total pixel No. = 64

Fig. 4. Two-dimensional optical neural network.

IV. Optical Implementation

The 2-D and 3-D parallel processing capabilities of the optical system make it an important candidate in solving the complex interconnections of a neural network.[5,6,8]

An optical neural network[6] is used for implementing the IPA model. The optical architecture is illustrated in Fig. 4. In this system, a high resolution video monitor is used to display the IWM, which also provides the source of incoherent light for processing. This system differs from the matrix–vector processor, as proposed by Farhat et al,[5] in which the positions of the input array and weighting matrix mask have been exchanged. Notice that this arrangement makes it possible to use a video monitor for IWM instead of a low resolution, low contrast SLM. A lenslet array consisting of $N \times N$ lenses is used to establish the optical interconnections. A moderate sized programmable SLM of $N \times N$ binary pixels serves as the input device. The light beam emanating from each block of the TV screen (i.e., submatrix of IWM) is passed through a specific lens of the lenslet array and is imaged onto the input SLM. The beams passing the SLM are then imaged onto the output plane by an imaging lens to form an $N \times N$ array output image which represents the product of a 4-D IWM with a 2-D input array. The output signals can be picked up by an $N \times N$ array photodetector.

To construct a closed loop neural network, the output signals from the detector array are fed back to the input SLM via a threshold circuit. By using the computer to calculate and modify the IWM, the optical neural network can be made both adaptive and programmable.

In experiments, Hopfield and IPA models are chosen to perform pattern recognition using noisy input patterns. B, P, and R are three letters stored in the weight matrix as shown in Fig. 5(a). The positive and negative parts of the IWMs for the IPA model are shown in Figs. 5(b) and (c), while those for the Hopfield model are displayed in Figs. 5(d) and (e). By comparing these two IWM sets it can be seen that the IPA model has two major advantages over the Hopfield model, namely, (1) fewer interconnections and (2) fewer gray levels. The latter is significant because the
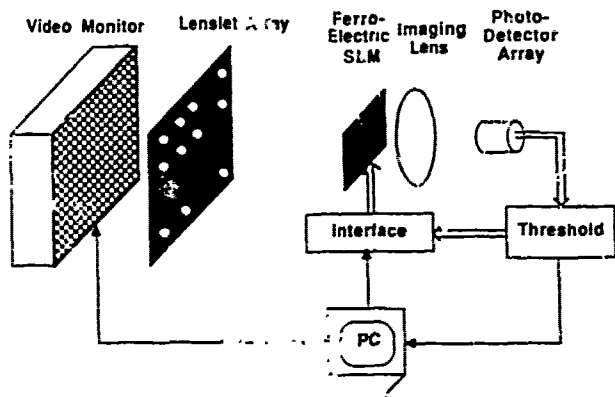

(a)

(b)   (c)

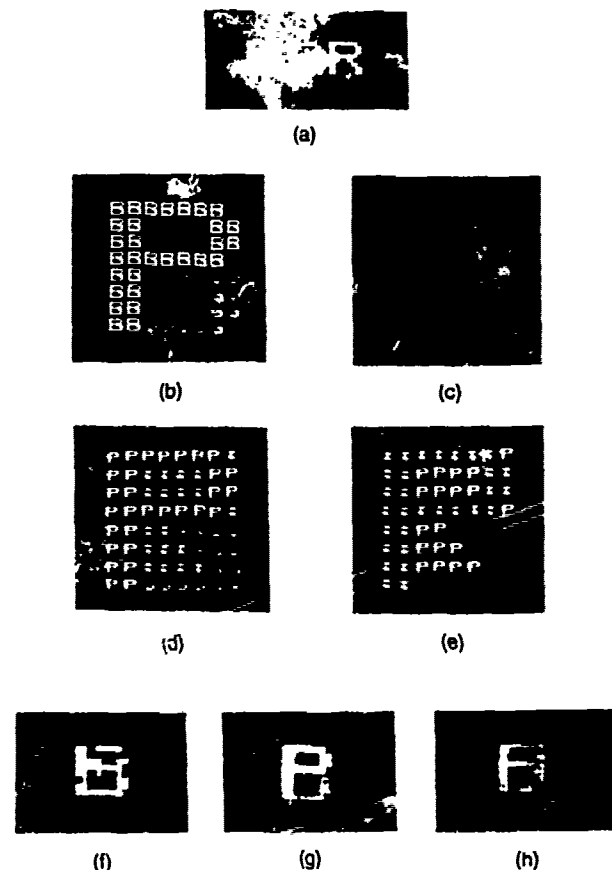(d)   (e)

(f)   (g)   (h)

Fig. 5. Experiments on an optical neurocomputer. Three similar English letters as reference patterns (a), positive and negative parts of the IWMs of the IPA model (b) and (c) and the Hopfield model (d) and (e), input pattern, SNR = 7 dB (f); results of pattern reconstruction by using IPA model (g) and Hopfield model (h).

IPA model requires only three gray levels to represent the IWM, whereas the Hopfield model needs $2M + 1$ gray levels, where M denotes the number of stored reference patterns.

The experimental results of these two models are obtained based on an input pattern B embedded in 30% random noise (SNR = 7 dB), as shown in Fig. 5(f).

The output patterns of the IPA and Hopfield models are shown in Figs. 5(g) and (h), respectively. Because of the curvature of the video monitor screen, the output results are somewhat distorted. Nevertheless, the results from the IPA model have been shown to be better than those from the Hopfield model.

V. Conclusions

We have proposed a new model—an IPA neural network. By using a simple logical rule, the common and special features of the reference patterns can be obtained, and the positive or negative interconnections can be assigned to each neuron. The IWM can be easily formulated, which requires merely an SLM with three gray levels to realize the IWM. Computer simulations and experiments reveal that the IPA model is more effective in performing pattern recognition (among similar patterns) than the models that deal only with the intrapattern association.

References

1. D. E. Rumelhart and J. L. McClelland, Eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vols. 1 and 2 (MIT Press, Cambridge, 1986).
2. T. Kohonen, *Self-Organization and Associative Memory* (Springer-Verlag, Berlin, 1984).
3. K. Fukushima, "A Neural Network for Visual Pattern Recognition," Computer 21, No. 3, 65–75 (1988).
4. G. A. Carpenter, and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," *Comput. Vision Graphics Image Process.* 37, 54–115 (1987).
5. N. H. Farhat, and D. Psaltis, "Optical Implementation of Associative Memory Based on Models of Neural Networks," *Optical Signal Processing*, J. L. Horner, Ed. (Academic, New York, 1987), pp. 129–162.
6. T. Lu, S. Wu, X. Xu, and F. T. S. Yu, "Two-Dimensional Programmable Optical Neural Network," Appl. Opt. 28, 4908–4913 (1989).
7. J. J. Hopfield, "Neural Network and Physical System with Emergent Collective Computational Abilities," Proc. Natl. Acad. Sci. USA 79, 2554–2558 (1982).
8. R. A. Athale. H. H. Szu, and C. B. Friedlander, "Optical Implementation of Associative Memory with Controlled Nonlinearity in the Correlation Domain," Opt. Lett. 11, 482–484 (1986).

APPENDIX 10.15

LCTV Neural Network

Optical neural network with pocket-sized liquid-crystal televisions

Francis T. S. Yu, Taiwei Lu, and Xiangyang Yang

Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

Don A. Gregory

U.S. Army Missile Command, Redstone Arsenal, Alabama 35898

A compact optical neural network that uses high-resolution liquid-crystal televisions has been constructed. System design considerations and an experimental demonstration of the liquid-crystal television neural network are reported.

Two of the most important features of an artificial neural network must be the possibility of massive interconnections and parallel-processing operations.[1,2] Recently, the use of optical techniques to perform the massive interconnection in a two-dimensional optical neural network was suggested by Farhat and Psaltis,[3] for which they used the basic concept of a vector–matrix optical processor. To provide the network with self-organizing and learning capabilities, a high-resolution, large-dynamic-range, programmable spatial light modulator is required for the interconnection weight matrix (IWM) construction. However, currently available spatial light modulators have small space–bandwidth products with limited gray levels, making it difficult actually to build such an optical neural network.

We recently proposed an optical architecture using a video monitor as a programmable IWM for the implementation of optical neural networks.[4,5] The resolution of the video monitor is approximately 1000 × 1000 pixels, and the dynamic range is approximately 256 gray levels. In fact, an 8 × 8 neuron optical neural network was built recently in our laboratory. However, the physical size of a video monitor makes the system large, and the curvature of the monitor screen also poses some problems. To alleviate these two major problems, in this Letter we present a compact optical neural network using pocket-sized liquid-crystal televisions (LCTV's).

Generally, an N-neuron network requires up to N^2 interconnections. As Hopfield suggested, the state of each neuron may be expressed by the following iterative equation:

$$v_i(n + 1) = f\left[\sum_{j=1}^{N} T_{ij}v_j(n)\right], \qquad i = 1, 2, \ldots, N, \quad (1)$$

where v_i is the state of the ith neuron, n is for the nth iteration, T_{ij} represents the connection strength between the jth and the ith neurons, and f is a nonlinear operator of the neurons.

As an example, if the neurons operate in bipolar states, f can be described as

$$f(x) = \begin{cases} 1 & x \geq t \\ -1 & x < t \end{cases}, \quad (2)$$

where t is the thresholding value. Note that the functional operation of Eq. (1) is a nonlinear operator that depends on the matrix–vector product, which can be expressed in a one-dimensional vector representation, such as

$$V(n + 1) = f[TV(n)], \quad (3)$$

where V is the state vector of the neurons and T is the IWM, which is also called the associative memory of the neural network.

It is apparent that, for a two-dimensional $N \times N$ neuron network, the iterative equation can be extended in the form

$$v_{lk}(n + 1) = f\left[\sum_{i=1}^{N}\sum_{j=1}^{N} T_{lkij}v_{ij}(n)\right], \quad (4)$$

where v_{lk} represents the state of the lkth neuron in an $N \times N$ space and T_{lkij} is a four-dimensional interconnection weight matrix. Note that the matrix T can be partitioned into an array of two-dimensional submatrices $T_{11ij}, T_{12ij}, \ldots T_{NNij}$, in which each submatrix is an $N \times N$ size.[3,6] Thus we see that a four-dimensional associative memory matrix can be displayed as an $N^2 \times N^2$ two-dimensional array representation (Fig. 1).

In recent developments of the LCTV, the imaging quality has been reported to be close to that of the commercially available high-resolution video monitors. For instance, the contrast ratio of a recent LCTV, with built-in thin-film transistors, can be higher than 30:1, and the dynamic range is approximately 16 gray levels, which can be continuously adjusted.[7] The optical architecture that we are presenting includes a Hitachi Model C5-LC1 5 in. (12.7-cm) color LCTV that is used to display the IWM and a Seiko Model LVD-202 2.7-in. (6.9-cm) color LCTV that is used to display the input patterns. The resolutions of
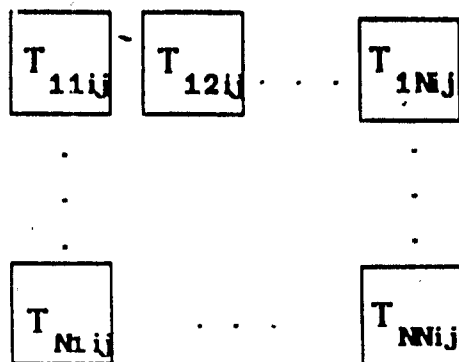
Fig. 1. Partition of a four-dimensional matrix T_{lkij} into an array of two-dimensional submatrices T_{11ij}, T_{12ij}, ... T_{NNij}.
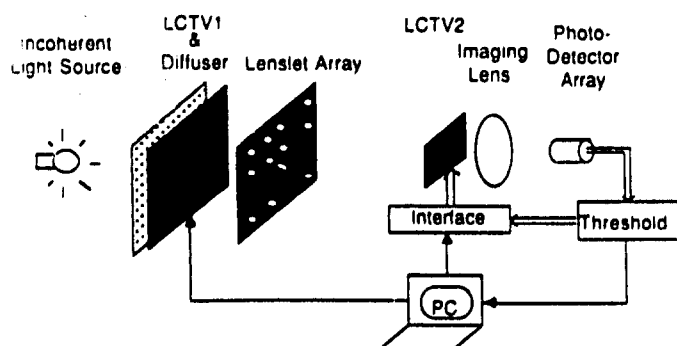

Fig. 2. Schematic of the compact optical neural network.

these LCTV's are 256 × 420 pixels and 220 × 330 pixels, respectively.

A diagram of a compact LCTV optical neural network is shown in Fig. 2. A neural network with 8 × 8 = 64 fully interconnected neurons was recently built in our laboratory, for which an 80-W xenon-arc lamp was used as the incoherent light source. In the present investigation the Hitachi LCTV (LCTV1) is used for the generation of the IWM, which consists of an 8 × 8 array of submatrices, with each submatrix having 8 × 8 elements. This IWM is then displayed on a fine diffuser immediately behind LCTV1. The Seiko LCTV (LCTV2) is used as an input device for the generation of the input patterns. The lenslet array, which consists of 8 × 8 lenses, provides the interconnections between the IWM and the input pattern. Each lens of the lenslet array images each of the IWM submatrices onto the input LCTV2 to establish the proper interconnections. Thus we see that the input matrix (i.e., LCTV2) is superimposed onto and multiplied (or interconnected) by all the IWM submatrices. This represents the connective part of Eq. (1), i.e.,

$$\sum_{l=1}^{N} \sum_{j=1}^{N} T_{lkij} v_{ij}(n).$$

The output result from LCTV2 is then imaged onto a charge-coupled-device (CCD) array detector. The signals collected by the CCD camera are then sent to a thresholding circuit, and the final results can be fed

back to the LCTV2 for a second iteration. Note that the data flow in the optical system is controlled primarily by a microcomputer (PC). For instance, the IWM's and the input patterns can be written onto the LCTV1 and the LCTV2 through the PC, and the PC can also make decisions based on the output results of the neural network. Thus the proposed LCTV optical neural network is indeed a programmable and adaptive neural network.
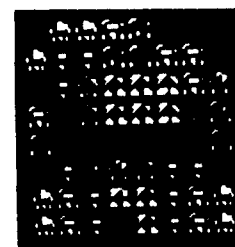
In the experimental setup of the compact optical neural network, an 80-W xenon-arc lamp is used as the light source. LCTV1 is used to display the IWM. An 8 × 8 lenslet array of plano-convex lenses (6-mm diameter, 72-mm focal length) sandwiched between two optical flat glass plates is used to establish the optical interconnection between LCTV1 and LCTV2; LCTV2 is used for input pattern generation. An imaging lens set (two identical plano-convex lenses of 100-mm diameter and 125-mm focal length, sealed face to face in a mount, with an effective 70-mm focal length) is used to image the output result from LCTV2 onto the CCD camera. This arrangement of plano-convex lenses has
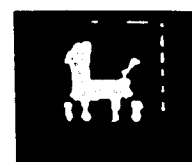

(a)
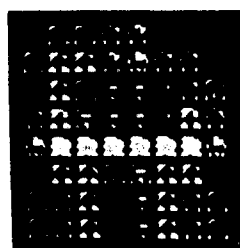
(b)

(c)

(d)

(e)

(f)

(g)

Fig. 3. Experimental results of the compact optical neural network. (a) Four cartoon images, a tree, a dog, a house, and an aircraft, used as reference patterns. (b) Positive and (c) negative parts of the IWM generated by using the interpattern association model. (d) Input pattern embedded in 25% noise, and (e) its output result: a dog, obtained in one iteration. (f) Input pattern in 25% noise, and (g) its output result: a tree, obtained in one iteration.

the ability to correct lens distortion. At the end of the system a Sony CCD camera (Model XC-77) is used as an output array detector.

The size of the optical neural network is approximately 65 cm × 12 cm × 12 cm. and the size of LCTV1 is approximately 12 cm. If a smaller LCTV is used for IWM generation (e.g., a 7-cm Epson color LCTV), then the system size can be further reduced to approximately 20 cm × 7 cm × 7 cm. In this case, the diameter of the lenses in a 16 × 20 lenslet array is within the submillimeter range, for which the lenslet array can be synthesized by using binary optics[8] or holographic techniques.

Since the LCTV1 has a resolution of 256 × 420 pixels. a maximum number of 16 × 20 = 320 neurons can actually be built. The matrix–vector operations of the proposed optical neural network are performed in parallel. The processing speed is limited by the speed of the LCTV's, which have an addressing speed of 30 frames/sec. Thus the operational speed of the proposed neural network with 8 × 8 neurons is approximately $30(8 \times 8)^2 = 1.2 \times 10^5$ interconnection operations per second. However, for a 16 × 20 neural network the operational speed can be as high as $30(16 \times 20)^2 = 3 \times 10^6$ interconnection operations per second.

In experiments, we have used the Hopfield, orthogonal projection, and interpattern association[9] models for the generation of the IWM's. One of the interpattern association experimental demonstrations is shown in Fig. 3. Four cartoon images, a tree, a dog, a house, and an aircraft, are used as reference patterns, and each letter is displayed in an 8 × 8 pixel array matrix [Fig. 3(a)]. We have used the interpattern association model (which is based on the association among the reference patterns[9]) for the construction of the IWM. The positive and negative parts of the IWM generated by the interpattern association model are shown in Figs. 3(b) and 3(c), respectively. There are, however, ways to represent the negative values in the IWM. by adding a bias into the IWM or by using an area-encoding method. However, for simplicity, the positive and negative parts of the IWM are sequentially displayed on the LCTV1 in our experimental demonstrations. An input pattern, i.e., a dog embedded in 25% random noise, is presented on LCTV2 [Fig. 3(d)]. The input pattern is then multiplied by the submatrices of the IWM. using the lenslet array. The positive and nega-

tive parts of the output results are captured by the CCD camera. and the subtraction and thresholding operations are performed by the PC, for which the output result is shown in Fig. 3(e). Thus a recognizable dog is reconstructed with only one iteration in this experiment. Similarly, when a tree with 25% noise was presented on LCTV1 [Fig. 3(f)], an output pattern of tree without any noise is obtained. as illustrated in Fig. 3(g). From this example, we can see that the compact optical neural network is capable of performing pattern-recognition and image-reconstruction operations.

In conclusion, we have shown that a compact optical neural network that uses high-resolution LCTV's can be constructed. The processing speed of this neural network can be as high as 3×10^6 interconnection operations per second, if a 16 × 20 lenslet array is used. To illustrate the performance of the LCTV optical neural network, an experimental result obtained with the interpretation association model is presented. Finally, we stress that low-cost, compact optical neural networks can indeed be built using inexpensive LCTV's.

References

1. D. E. Rumelhart and J. L. McClelland, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, Mass., 1986), Vols. 1 and 2.
2. J. J. Hopfield. Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
3. N. H. Farhat and D. Psaltis, in *Optical Signal Processing*, J. L. Horner. ed. (Academic. Orlando. Fla.. 1987), p. 129.
4. S. Wu, T. Lu, X. Xu, and F. T. S. Yu. Microwave Opt. Technol. Lett. **2**, 252 (1989).
5. T. Lu, S. Wu, X. Xu, and F. T. S. Yu. Appl. Opt. **28**, 4908 (1989).
6. N. H. Farhat. D. Psaltis. A. Prata. and E. Faek. Appl. Opt. **24**, 1469 (1985).
7. H. K. Liu and T. H. Chao. Appl. Opt. **28**, 4722 (1989).
8. W. Veldkamp, in *Digest of First Annual International Meeting on Neural Networks* (Pergamon. New York. 1988).
9. T. Lu, X. Xu. S. Wu. and F. T. S. Yu, Appl. Opt. **29**, 284 (1990).

APPENDIX 10.16

Nonconventional JTC

Nonconventional joint-transform correlator

Francis T. S. Yu, Chenhua Zhang, and Yong Jin

Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

Don A. Gregory

U.S. Army Missile Command, Redstone Arsenal, Alabama 35898

A nonconventional joint-transform correlator (NJTC) is discussed. We show that the shift-invariant property of the usual joint-transform arrangement can be preserved. The advantages of the NJTC are the efficient use of the light source, the use of smaller transform lenses, higher correlation peaks, and a higher carrier fringe frequency.

A conventional joint-transform correlator[1-3] (JTC) offers the advantages of the avoidance of normal matched-filter synthesis, a higher space–bandwidth product, a lower carrier frequency, a higher degree of modulation, and suitability for real-time implementation. However, the conventional JTC also suffers several major drawbacks, e.g., the inefficient use of illumination light, the requirement of a larger transform lens, a more stringent spatial coherent requirement, and the overall small size of the joint-transform spectrum. In this Letter we describe an alternative approach that alleviates some of these limitations.

The quasi-Fourier transformation of an optical system as defined here is shown in Fig. 1. We assume that the detection plane P_2 is located a small distance δ away from the back focal plane of the transform lens. With the usual Gaussian lens equation, the image of P_2 can be shown to be located at a distance

$$L = f^2/\delta \qquad (1)$$

away from input plane P_1. We now refer to the far-field (i.e., Fraunhofer) diffraction arrangement of Fig. 2. To achieve a Fourier diffraction pattern at plane P_2', it is required that $\Delta L'$ be sufficiently small compared with the illuminating wavelength λ, i.e.,

$$\Delta L' \leq \lambda/4. \qquad (2)$$

In view of the similar right triangles of Fig. 2, L' can be approximated by

$$L' \simeq \frac{(b/2)^2}{2\Delta L'}, \qquad (3)$$

where b is the size of the input object.

To obtain a Fourier diffraction pattern at the observation screen, it is required that the inequality of relation (2) hold. Thus the distance to the observation screen should be

$$L' \geq b^2/2\lambda. \qquad (4)$$

It is apparent that, to achieve a quasi-Fourier transformation at P_2 as depicted in Eq. (1), one should have

$$L \geq L'. \qquad (5)$$

Thus the allowable deviation from the focal length should be

$$\delta \leq 2\lambda(f/b)^2. \qquad (6)$$

Note that this result is known as the focal tolerance or the focal depth of the optical system.[4]

One of the most important aspects of an optical correlator must be the shift-invariant property of the Fourier transformation, from which the position of the object (i.e., the target) can be extracted. We now investigate the constraints in object translation under the quasi-Fourier-transform regime. Consider the quasi-joint-Fourier-transform configuration of Fig. 3. A displacement of the object spectrum at the quasi-Fourier domain P_2 is anticipated. Referring to the triangular configuration in Fig. 3, the amount of displacement along the vertical direction is

$$a \approx \frac{\delta}{f} d, \qquad (7)$$

where d is the translation of the object.

To ensure good overlapping object and reference spectra at the quasi-Fourier domain P_2, the amount of displacement should be small compared with the size of the main lobe of the object spectrum, i.e.,

$$a \leq \lambda f/2b. \qquad (8)$$

If we simply substitute Eq. (7) into relation (8), the amount of allowable object displacement is

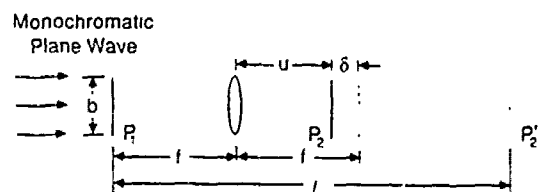$$d \leq \frac{f^2\lambda}{2b\delta}. \qquad (9)$$

Fig. 1. Quasi-Fourier transformation. P_2' is the image of P_2.
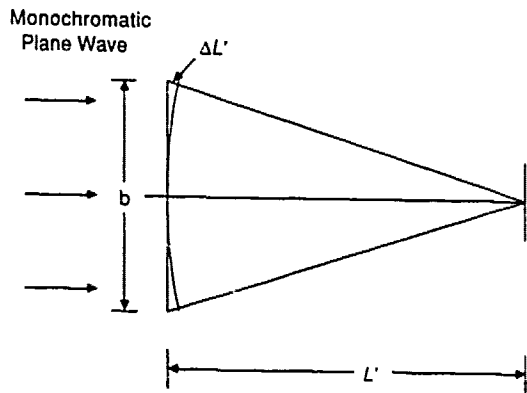

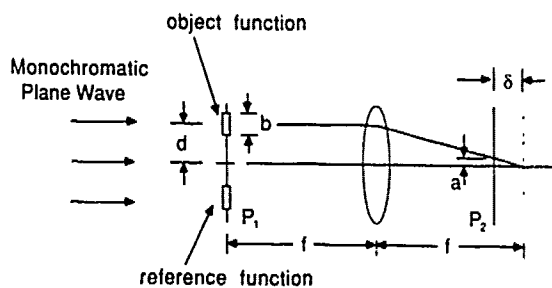
Fig. 2.   Fourier diffraction condition.

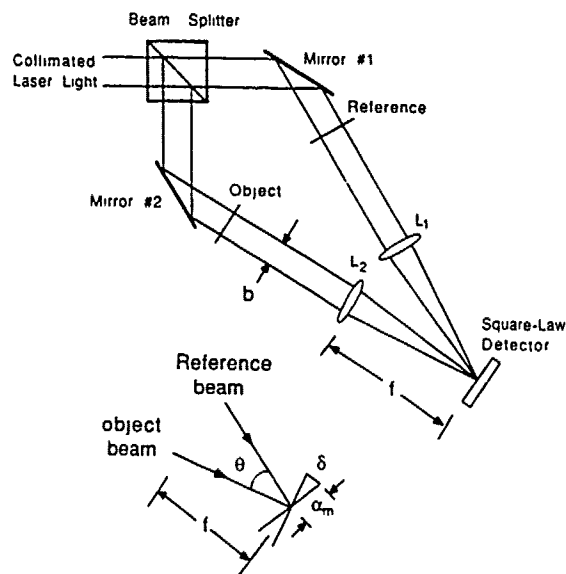Fig. 3.   Quasi-joint-Fourier transformation.

Fig. 4.   Nonconventional joint Fourier transformation.

To recapitulate, the separation between the object and the reference functions should be within the constraint of relation (9). Under this condition the object and reference spectra will be mostly overlapped, producing interferometric fringes (i.e., the carrier frequency) such that the shift-invariant property can be preserved.

We now describe the nonconventional joint-transform technique as illustrated in Fig. 4. Let us assume that $2\alpha_m$ is the size of the object spectrum, with α_m written as

$$\alpha_m = f\lambda f_{max} = f\lambda W/b, \qquad (10)$$

where f_{max} is the upper spatial frequency content of the object and $W = bf_{max}$ is the space–bandwidth product of the object. Note that the incident angle at the square-law detector, formed by the object and the reference beams, can be written as

$$\theta \simeq \arctan(\delta/\alpha_m), \qquad (11)$$

where δ is the maximum focal depth of relation (6). By substituting relation (6) and Eq. (10) into relation (11), the constraint of the incident angle is

$$\theta \leq \arctan(2f/bW), \qquad (12)$$

where f is the focal length of the transform lens, b is the size of the object, and W is the space–bandwidth product of the object. To have an idea of the magnitudes involved, we assume that the focal length $f = 500$ mm, the object size $b = 10$ mm, and the space–bandwidth product of the object $W = 200$; thus the incident angle must be $\theta \leq 27°$.

We further note that the proposed nonconventional joint Fourier transformation may also be achieved with a different arrangement, as shown in Fig. 5, in which we assume that the constraints of the focal

$$\delta \leq 2\lambda\left(\frac{1}{b}\right)^2$$

$$\theta \leq \arctan\left(\frac{2f}{bW}\right)$$
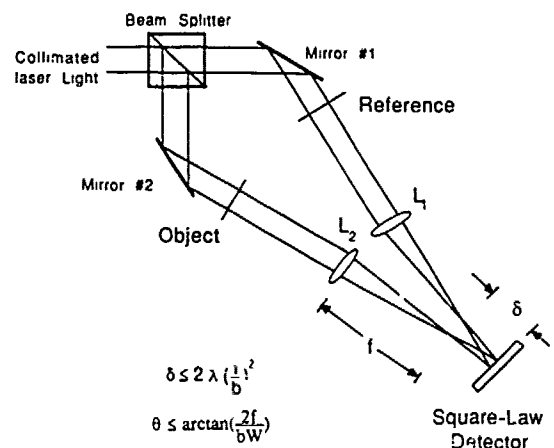
Fig. 5.   Nonconventional quasi-joint-Fourier transformation.

Input Object   Reference

(a)

(b)

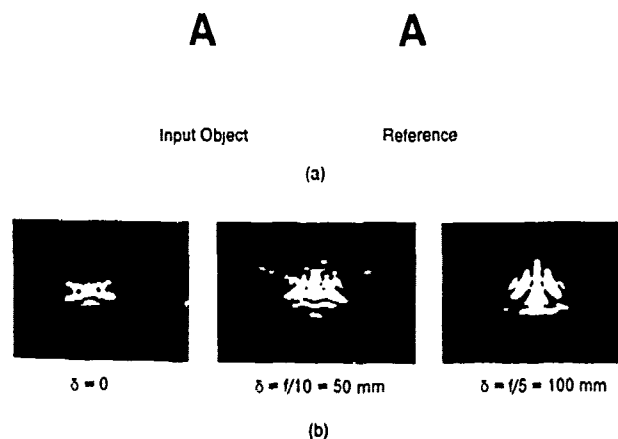Fig. 6.   (a) Input object and reference function.   (b) Output correlation spots.

A   **A**

Input Object   Reference

(a)

δ = 0   δ = f/10 = 50 mm   δ = f/5 = 100 mm

(b)

Fig. 7. (a) Input and reference objects. (b) Joint-transform holograms.

δ = 0   δ = f/10 = 50 mm   δ = f/5 = 100 mm

(a)
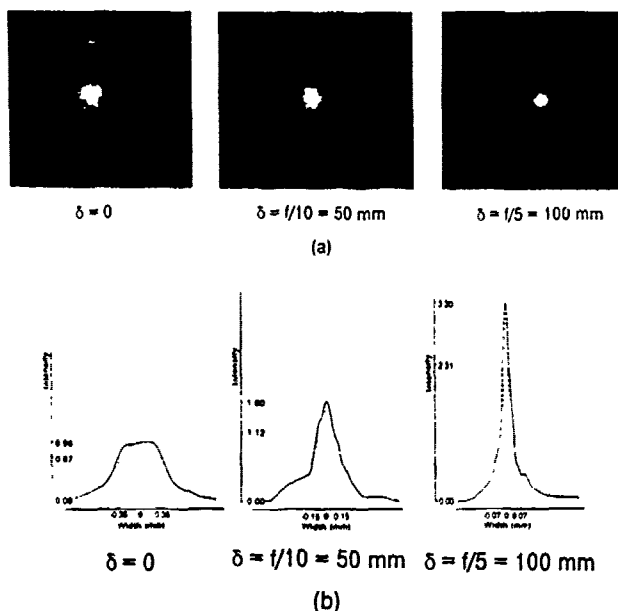

δ = 0   δ = f/10 = 50 mm   δ = f/5 = 100 mm

(b)

Fig. 8. (a) Output correlation spots. (b) Photometer traces of the output correlation peaks.

depth of relation (6) and the allowable incident angle of relation (12) hold. These optical setups have the advantage of producing a larger joint-transform spectrum, for which a higher read-out correlation peak can be obtained. Furthermore, if the target displacement is within the shift-invariant constraint of relation (9), the proposed nonconventional JTC (NJTC) can also extract the spatial location of the target.

To illustrate the shift-invariant property of the NJTC, an input transparency with four English characters and a reference letter [Fig. 6(a)] was used. The NJTC structure used in this experiment is shown in Fig. 4. The joint-transform power spectrum was recorded by a photographic transparency, which is called a joint-transform hologram (JTH). By simply illuminating the recorded JTH with coherent light, cross correlations of input characters with the reference letter can be viewed at the output plane shown in

Fig. 6(b). Two autocorrelation peaks, representing the locations of the input character G, can be observed. Thus we see that the shift-invariant property of the NJTC is preserved. The incident angle for this experiment was set at ~10°.

To illustrate the results obtained using the quasi-Fourier-transformation arrangement of Fig. 5, the input object and reference function of Fig. 7(a) were used. In this experiment three JTH's for $\delta = 0$, $\delta = f/10 = 50$ mm, and $\delta = f/5 = 100$ mm were recorded [Fig. 7(b)]. Throughout the making of the set of holograms, the incident angle was maintained at $\theta \simeq 10°$. Note that the scale of the joint power spectrum enlarges as δ increases. Again, by simple coherent illumination of the JTH's, their output correlation peaks were obtained [Fig. 8(a)]. Photometer traces of the corresponding correlation peaks are shown in Fig. 8(b). From these results we see that the peak intensity increases as δ increases, while the size of the correlation spot decreases as δ increases. This phenomenon is primarily due to the size of the joint-transform spectrum, as seen in Fig. 7(b). Thus the proposed NJTC has the ability to improve the output signal-to-noise ratio and the accuracy of the signal detection.

Furthermore, with a larger incident angle θ, it is also possible to increase the spatial carrier frequency of the JTH. This would allow a larger separation between correlation spots away from the zeroth-order diffraction. However, higher spatial carrier fringes would impose a higher resolution constraint on the square-law detector used.

We have presented a NJTC technique that alleviates the inefficient use of the light source as in the conventional JTC. We have shown that with a quasi-Fourier arrangement the shift-invariant property of a NJTC can be maintained. The major advantages of the NJTC are the efficient use of incident light, the use of smaller transform lenses, higher correlation peaks, and a higher carrier fringe frequency.

Since a joint-transform processor[5] can, in principle, process all the information that a conventional coherent optical processor can, we believe that the NJTC possesses many capabilities beyond those discussed here.

References

1. C. S. Weaver and J. W. Goodman, Appl. Opt. 5, 1248 (1966).
2. F. T. S. Yu and X. J. Lu, Opt. Commun. 52, 10 (1984).
3. D. A. Gregory, J. A. Loudin, and F. T. S. Yu, "Illumination dependence of the joint-transform correlation," Appl. Opt. (to be published).
4. M. Born and E. Wolf, *Principle of Optics*, 2nd ed. (Pergamon, New York, 1964), p. 441.
5. F. T. S. Yu and J. E. Ludman, "Joint Fourier transform processor," Microwave Opt. Technol. Lett. 1, 374 (1988).

APPENDIX 10._7

Multi-Target Tracking

Data association multiple target tracking using a phase-mostly liquid crystal television

Eddy C. Tam, MEMBER SPIE
University of North Carolina at Charlotte
Department of Physics
Charlotte, North Carolina 28223

Francis T. S. Yu, FELLOW SPIE
Aris Tanone, MEMBER SPIE
The Pennsylvania State University
Department of Electrical Engineering
University Park, Pennsylvania 16802

Don A. Gregory
U.S. Army Missile Command
Research Directorate
Redstone Arsenal, Alabama 35898-5248

Richard D. Juday, MEMBER SPIE
NASA Johnson Space Center
Houston, Texas 77058

Abstract. This paper presents a technique of using data association target tracking in a motion sequence via an adaptive joint transform correlator. The massive data in the field of view can be reduced to a few correlation peaks. The average velocity of a target during the tracking cycle is then determined from the location of the correlation peak. We have used a data association algorithm for the analysis of these correlation signals, with which multiple targets can be tracked. A phase mostly LCTV is used in the hybrid joint transform correlation system, and simultaneous tracking of three targets is demonstrated.

Subject terms: optical pattern recognition; target tracking; liquid crystal television; joint transform correlators.

Optical Engineering 29(9), 1114-1121 (September 1990)

CONTENTS

1. INTRODUCTION

One of the major advantages of optical signal processing is the parallelism in handling high space-bandwidth signals. However, the price paid is the complexity and stringent alignment of an optical system. On the other hand, its digital counterpart offers flexibility and programmability, while sacrificing full parallelism. In this paper, we describe a hybrid optical technique to tackle a multitarget tracking problem. Let us assume that there are multiple targets whose dynamic states (e.g., locations, velocities, etc.) are to be evaluated in a motion sequence. At every frame a new scene is captured by an optical sensor. The space-bandwidth product in each frame is so high that it is extremely difficult for an all electronic computer to process on a real-time

basis. Although parallel processing can be performed by an optical system, it is not feasible to build an optical system that performs precise decision making. The hybrid optical system that we propose uses a joint transform correlator to carry out the correlation operation and a data association algorithm for multiple target tracking.

Target tracking with an optical correlator was first demonstrated by VanderLugt in 1974.[1] Subsequently, real-time tracking using a VanderLugt correlator has been reported.[2,3] The development of a joint transform correlator (JTC),[4,5] on the other hand, offers an alternative for target tracking.[6] Simplicity, easier alignment, and the avoidance of matched filter synthesis are the major advantages of using a JTC. It also has the capability to adapt to new input scenes: in the proposed architecture, the previous frame is used as a reference image to correlate with the target in the current frame. Assuming that the processing cycle of the system is sufficiently short, then a target will look much the same after a few sequential frames, and hence strong correlation signals will be generated at the output plane. The position measurements of the correlation peaks can be interpreted as the average velocities of the targets within a tracking system. Based on these measurements, the dynamic states of the targets can be updated. Since multiple correlation peaks are generated in a multitarget tracking problem, a data association algorithm can be employed to associate each of the peaks to the correct paths of the targets' motion. It must be mentioned that the adaptive property of this system enables the tracking of targets under change in orientation, scale, and perspective during the course of their movement.

In this paper we demonstrate the use of a hybrid JTC based on the aforementioned technique of multiple target tracking. A liquid crystal television was used in our experiment, and the implementation of this spatial light modulator (SLM) using phase-mostly modulation mode is discussed in detail. To simplify the subject matter of this paper, the problem of initial target acquisition is not discussed here.

2. TRACKING ALGORITHM

Let $f^n(x,y)$, $n = 1$ to N, be the targets in an image sequence, each of which are moving independently. Two sequential scenes are displayed on the input plane of a JTC such that the previous frame $(k - 1)$ and the current frame (k) are positioned in the upper and lower halves of the LCTV, respectively. The corresponding image function is given by

$$T(x,y) = \sum_{n=1}^{N} f_{k-1}^n \left(x - x_{k-1}^n, y - y_{k-1}^n - \frac{a}{2} \right)$$
$$- \sum_{n=1}^{N} f_k^n \left(x - x_{k-1}^n - \delta x_k^n, y - y_{k-1}^n - \delta y_k^n + \frac{a}{2} \right) . \quad (1)$$

where a is the height of the active aperture of the LCTV, (x_{k-1}^n, y_{k-1}^n) are the locations of the targets in the $k - 1$ frame, and $(\delta x_k^n, \delta y_k^n)$ are the relative translations of the targets from the $k - 1$ frame to the k frame. The corresponding joint transform power spectrum (JTPS) is then introduced as modulation at the input plane for the second Fourier transformation. It can be shown that the output complex light distribution is given by

$$C(x,y) =$$

$$\sum_{n=1}^{N} \sum_{m=1}^{N} R_{k-1,k-1}^{m,n}(x + x_{k-1}^m - x_{k-1}^n, y + y_{k-1}^m - y_{k-1}^n)$$

$$- \sum_{n=1}^{N} \sum_{m=1}^{N} R_{k,k}^{m,n}(x + x_{k-1}^m - \delta x_k^m - x_{k-1}^n + \delta x_k^n, y + y_{k-1}^m - \delta y_k^m - y_{k-1}^n + \delta y_k^n)$$

$$- \sum_{n=1}^{N} \sum_{m=1}^{N} R_{k-1,k}^{m,n}(x + x_{k-1}^m - x_{k-1}^n + \delta x_k^n, y + y_{k-1}^m - y_{k-1}^n + \delta y_k^n + a)$$

$$- \sum_{n=1}^{N} \sum_{m=1}^{N} R_{k,k-1}^{m,n}(x + x_{k-1}^m - \delta x_k^m - x_{k-1}^n, y + y_{k-1}^m - \delta y_k^m - y_{k-1}^n - a) . \quad (2)$$

where

$$R_{k,k}^{m,n}(x - \alpha, y - \beta) = f_k^m(x,y) \otimes f_k^n(x,y) \star \delta(x - \alpha, y - \beta) \quad (3)$$

is the correlation function between f_k^m and f_k^n located at (α, β), and $\otimes$ and $\star$ denote the correlation and convolution operations, respectively.

Note that the autocorrelation functions $R_{k,k}^{n,n}$ and $R_{k-1,k-1}^{n,n}$ are located at the origin of the output plane. The correlation between the same target in the k frame and the $k - 1$ frame, given as, $R_{k,k-1}^{n,n}$, is diffracted around $(\delta x_k^n, \delta y_k^n + a)$ and $(-\delta x_k^n, -\delta y_k^n - a)$. The rest of the terms ($R_{k,k-1}^{n,n}$ and $R_{k,k}^{m,n}$ for $m \neq n$) represent the cross-correlation functions between target m and target n. It must be stated that in this multiple target tracking problem, we have assumed that the targets in any single input frame do not look alike, so the correlation functions $R_{k,k-1}^{m,n}$ and $R_{k,k}^{m,n}$ generate much weaker correlation signals than that generated by the function $R_{k,k-1}^{n,n}$. Furthermore, the size and the gray level of the targets are also assumed to be relatively congruous with each other, so the intensities of the correlation peaks generated from $R_{k,k-1}^{n,n}$, $n = 1$ to N, are not significantly different.

By translating the coordinate origin from the optical axis to $(0, a)$, the correlation peaks that are generated from the same target in the $k - 1$ and the k frame are then located at $(\delta x_k^n, \delta y_k^n)$ in this new coordinate system. If these correlation peaks are associated with the proper target motions, then the locations of the targets in the k frame are given by

$$x_k^n = x_{k-1}^n + \delta x_k^n, \quad y_k^n = y_{k-1}^n + \delta y_k^n, \quad n = 1 \text{ to } N . \quad (4)$$

Equation (4) shows that the new locations of the targets in the k frame can be updated based on their locations in the previous frame and the position measurements of the correlation peaks. It is apparent that the k and $k + 1$ frames can be sent to the input plane in the next tracking cycle, so multiple targets can be tracked simultaneously on a near real-time basis.

Note that the position of the correlation peak in the new coordinate system represents the average velocity of a target during the sampling interval δt, i.e.,

$$\bar{x} = \delta x/\delta t, \quad \bar{y} = \delta y/\delta t . \quad (5)$$

Therefore, with a constant sampling interval and assuming that the sensor is not slewing or panning, the new coordinate system represents the plane of average velocity. For example, targets moving with constant velocity produce stationary correlation peaks, targets moving with constant acceleration generate correlation peaks that are located at equally separated intervals on a straight line, and correlation peaks located at the origin of the new coordinate plane correspond to stationary objects or background scene. In our experimental demonstration, multiple targets are traveling at a different velocity and in different directions, thus the correlation peaks generated are located at different positions at the velocity plane.

3. DATA ASSOCIATION

In general, the motion of the targets can be represented by dynamic models, which are governed by well-known laws of physics. Unpredictable changes in target motions, commonly called maneuvers, can also be treated as gradual changes of motion parameters, if the sampling interval is sufficiently short compared with the maneuver time. Therefore, it is not difficult to associate the measurements in the velocity plane with the targets, based on the past history of the targets. Any prior knowledge of the targets before tracking will also be helpful in this association process. For example, if one target is known to travel much faster than the others, then its correlation peak is more likely to be located farther away from the origin in the velocity plane than those of slower ones.

In this section, a parameter-based data association algorithm using a Kalman filtering model is described. It is assumed that the sampling interval δt is sufficiently short that the dynamic parameters (velocity, acceleration, angular velocity, etc.) of the targets are fairly consistent within a few sequential frames. Thus, given the dynamic parameters of a target in the k frame, the parameters in the next frame should be rather predictable. Let $z(k)$ be the measurement at the k frame and $\hat{z}(k \mid k - 1)$ be the predicted measurement in the k frame, based upon the information evaluated up to the $k - 1$ frame. Then, the innovation (or measurement residual), defined as

$$v(k) = z(k) - \hat{z}(k \mid k-1) . \qquad (6)$$

can be used to evaluate the likelihood of association between $z(k)$ and the target under consideration. In stochastic models, one would evaluate the normalized squared distance D from the measured $z(k)$ to the current track,

$$D = v'S^{-1}v , \qquad (7)$$

where v' is the transpose of v and S is the innovation covariance of the target, which is basically the variance of the estimated states. The values of $\hat{z}(k \mid k - 1)$, S, and the like can be evaluated by applying a standard Kalman filter to the dynamic model. The use of Kalman filtering in stochastic modeling[7] is a well-known subject on its own and will not be discussed here. However, a brief discussion will be given in Sec. 7, followed by an explanation of a simplified filtering process to illustrate the capability of this multiple target tracking technique.

A data association process can be carried out as follows:

Step 1: At the $k - 1$ frame, the dynamic parameters of the N targets are determined at track 1 to N.
Step 2: At the k frame, N new measurements are made, given as $a, b, ..., N$. The normalized square distances are then computed:

$$D_{1a} = v'_{1a}S_1^{-1}v_{1a}, \quad D_{1b} = v'_{1b}S_1^{-1}v_{1b}, ..., \quad \text{etc.} .$$

and similarly for $D_{2a}, D_{2b}, ..., D_{3a}, ...,$ and so on.
Step 3: The most likely association is given by choosing the possible combination of Ds that yields the minimum sum.

To initiate the tracker, all the measurements in the first few tracking cycles are used to set up multiple potential tracks. Tracks that have inconsistent dynamic parameters are dropped until only a single track is assigned to each target. Recall that the measurement in each cycle is the average velocity of a target during the sampling period. Therefore, dynamic parameters can be assigned quickly in a few cycles; only two or three cycles are needed to determine targets with constant velocity or constant acceleration. The tracking scheme discussed so far, however, does not determine the initial positions of the targets in the first frame nor identify the measurements to the targets due to the nature of this adaptive correlation scheme. Therefore, an initial acquisition scheme is needed to perform this task before the start of the adaptive tracker. This can be done by using prestored reference images located at fixed positions at the image plane. Further analysis of the initial target acquisition technique is beyond the scope of this paper.

4. DIFFRACTION EFFICIENCY AND RESOLUTION

In multiple object correlation, the JTPS is encoded with multiple carrier frequencies. The detection efficiency of the system is therefore inevitably lower when compared with that for single object correlation.[8] Consider when N identical objects, each located at x_n, $n = 1$ to N in a 1-D space, are the input objects to a JTC. The input function is then given as

$$T(x,v) = \sum_{n=1}^{N} f(x - x_n) \qquad (8)$$

If the JTPS is linearly recorded and displayed on a SLM, the normalized amplitude transmittance on the SLM becomes

$$I(u,v) = \frac{1}{N^2 \mid F(0,0) \mid^2} \mid F(u,v) \mid^2 \left[N + \sum_{n=1}^{N} \sum_{m>n}^{N} 2\cos(x_n - x_m)u \right] .$$

$$(9)$$

where the normalization constant is set to be the magnitude at $u = 0$. The correlation output is obtained by taking the inverse Fourier transform of $I(u,v)$, and the intensity of the correlation function at $x_n - x_m$ is given as

$$C(x,y) = \frac{1}{N^4 \mid F(0,0) \mid^4} \mid f(x,y) \otimes f(x,y) \star \delta(x - x_n + x_m) \mid^2 . \qquad (10)$$

The $1/N^4$ factor seems to suggest that a JTC is of little value for multiple object correlation. However, it is obvious that this factor comes in as the normalization constant in Eq. (9). In practical considerations, the energy at the origin is much higher than that in the other frequency components; it is therefore unrealistic to use this value as the normalization constant. In fact, in most cases the dynamic range of a power spectrum is so large that no detector can record the spectrum linearly. Therefore, imposing an upper limit on the power spectrum is unavoidable. Moreover, it can be seen that the bracketed terms in Eq. (9) are a summation of sinusoidal signals of different spatial frequencies. The bias level build-up in this summation significantly reduces the modulation of the fringes. However, due to the sensitivity of a practical detector, part of this bias will be eliminated, hence increasing the modulation of the fringes. Therefore, in most experiments, replacing the normalization process on the power spectrum with a nonlinear recording process bounded between the lower and upper threshold is unavoidable, and the number of objects in a JTC will affect the diffraction efficiency in a highly nonlinear fashion. If the upper and lower threshold are properly chosen such that the modulation of the fringes is amplified by this clipping effect, then a much better diffraction efficiency can be achieved even for multiple object correlation. The thresholding levels can be adjusted simply by changing the output power of the illuminating source, by software programming, or by changing the recording property of the square law detector.[9] In this experiment, the first two techniques mentioned were utilized to optimize the correlation output.

Due to the nature of this adaptive tracking technique, the correlation peaks generated from different targets with almost identical velocities will fuse together, hence reducing the resolution of the correlation spots. If W is the width of a target, then $2W$ is the minimum separation for the two correlation functions to be completely separated. However, due to the nonlinear thresholding effect imposed on the power spectrum, as discussed previously, the width of a correlation peak will be much sharper than $2W$. Nevertheless, the minimum separation for two peaks to be distinguishable must be greater than or equal to the full width half maximum (FWHM) of the peak. To resolve the measurements for two moving targets, their difference in velocity must satisfy the condition

$$(\Delta \bar{x}/FWHM_x)^2 + (\Delta \bar{y}/FWHM_y)^2 > 1 . \qquad (11)$$

If the sampling interval is too short compared with the velocities of the targets, the correlation peaks will be crowded together near the origin (of the translated coordinate) and will become unresolvable. Therefore, this resolution criterion imposes a limitation to this adaptive tracking algorithm, for which an optimum sampling rate should be used.

5. HYBRID MODULATION PROPERTY

Previous applications of LCTVs in optical processors mostly used the polarization modulation property of twisted nematic liquid crystals. Amplitude modulation can thus be obtained according to Malus's law. Even under binary phase-only modulation, as proposed previously,[10,11] the phase modulation actually results from the polarization modulation by orienting the analyzer to be perpendicular to the bisector between the ON and OFF states of the LCTV. However, besides the polarization modulation property of a 90° twisted nematic liquid crystal cell, there is also an inherent birefringence effect associated with the molecule. The birefringence is altered as the structure of the liquid crystal starts to deform when the applied electric field is above a certain threshold level (the twisted nematic liquid crystal deformation threshold). This results in phase-only modulation. When the applied field is sufficiently high and reaches another threshold value (i.e., the optical threshold), the twisted structure starts to untwist itself and polarization modulation takes place. In this region, the LCTV is modulated in a hybrid mode: both phase and polarization modulations. Therefore, phase-mostly modulation can be achieved by setting the bias voltage of the LCTV at the minimum value. This effect was documented by Blinov[12] and has been discussed recently by Konforti et. al.[13] Applications of this phase modulating property have also been reported.[14,15]

It must be mentioned that for optimum phase modulation effect, the polarization direction of the incident light must be aligned parallel to the molecular director on the LCTV's front surface. No phase modulation was observed in our experiment when the polarization director was perpendicular to the molecule director. This has also been mentioned in both Refs. 13 and 14. This effect can be explained as follows. When the polarization plane of the incident light is parallel to the director of the liquid crystal molecules, the electric field vector (E-vector) is parallel to the molecules; hence, the incidence light becomes extraordinary to the liquid crystal cell. As the electric field propagates through the twisted cell, its plane of polarization follows the twist such that the E-vector is always parallel to the molecules, and therefore phase modulation is achieved. However, if the polarization plane of the incident light is perpendicular to the molecule, then it is an ordinary light to the cell. As the electric field propagates through the twisted cell, its plane of polarization remains perpendicular to the molecule, and there is no relative phase shift. It is interesting to note that the original polarizer attached to the LCTV used in this experiment (Seiko LVD 202 color LCTV, 220×240 pixels) is found to be aligned perpendicular to the molecular director. We suspect that by doing so, the undesirable coloring effect due to birefringence can be eliminated, and proper color selection is solely controlled by internal circuitry driving the RGB pixels.

To measure the phase modulation property of this LCTV, we have removed both polarizer and analyzer from the liquid crystal display panel. A Mach-Zehnder interferometer similar to the one mentioned in Ref. 13 was used to measure the phase modulation property of the LCTV. The minimum bias voltage level ($V = 0.978$, measured across the brightness control switch) was set at the LCTV, and the different gray levels (from 0 to 255) generated from a Datacube frame grabber (AT428) were sent to the TV through a video/rf modulator. The interference fringes observed at the output plane were recorded. It was found that as the gray level varied, the fringes started to shift, and maximum
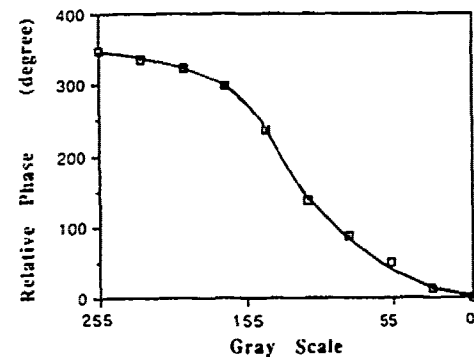

Fig. 1. Plot of the relative phase shift as a function of gray level. The gray level is labeled in descending order since the applied voltage increases as the input gray level decreases.

observed phase modulation was approximately 340°. The phase modulation angle versus input gray level is plotted in Fig. 1. We note that this LCTV is operated with negative logic: minimum input gray level (0 gray level) corresponds to maximum applied voltage across the cell and maximum gray level (255) corresponds to minimum applied voltage, and therefore we labeled the x-axis in descending order. The polarization modulation property of the LCTV also was measured at the same bias voltage, and it was found that polarization modulation occurs at gray levels below 60. Therefore, an input image with gray scale above 60 is phase-mostly modulated, and an image with gray scale below this value is encoded with both phase ϕ and polarization θ modulations.

To investigate the effect of this hybrid modulation in a JTC, consider the plane of principal polarization to be decomposed into two orthogonal components $\hat{\alpha}$ and $\hat{\beta}$. The complex light amplitude behind the LCTV would be

$$E_f(x,y) = |E|\exp[i\phi(x,y)][\hat{\alpha}\cos\theta(x,y) + \hat{\beta}\sin\theta(x,y)] , \qquad (12)$$

where $\phi(x,y)$ and $\theta(x,y)$ are the phase encoding and polarization encoding of an input function $f(x,y)$. If an analyzer is inserted behind the LCTV with its director along the $\hat{\alpha}$ direction, then only the $\hat{\alpha}$ component in Eq. (12) emerges through the LCTV. The image is then subjected to phase and amplitude modulations. However, if the analyzer is not inserted, then both the $\hat{\alpha}$ and $\hat{\beta}$ components will contribute at the Fourier plane. Since these two components are orthogonal, the Fourier transform spectra resulting from the first and second terms in Eq. (12) will add in terms of intensity rather then in complex amplitude. Moreover, the $\hat{\alpha}$ and $\hat{\beta}$ terms are basically two reversed contrast images; the power spectra of the two images are basically the same. It can be shown that if $f(x,y)$ is a binary image, then the two power spectra are exactly identical except for different dc magnitudes, according to Babinet's principle. Besides, it is the fringe modulation that is important in a JTC, and it can be shown that both the $\hat{\alpha}$ and $\hat{\beta}$ components produce the same fringe modulations. Therefore, a JTC operating in hybrid phase/polarization modulations (without an analyzer) would offer performance similar to a phase/amplitude coupled LCTV. The transmission of the LCTV, however, will be much higher due to the absence of the analyzer.

6. EXPERIMENTAL SETUP

Figure 2 shows the JTC architecture we have employed in the experiment. The current frame and the previous frame in a motion sequence were displayed onto the lower and upper halves of the LCTV. After the first Fourier transformation, the JTPS captured by CCD2 was then sent back to the LCTV for the second transformation. The output correlation was detected again by CCD2, with which the locations of the correlation peaks were determined by sequentially searching through the correlation plane. Finally, a data association algorithm was used to sort the measurements and produced the proper tracks for the targets.

To reduce the physical size of the correlator, we have used a negative lens behind the Fourier lens so that a very long effective focal length was obtained with a relatively short physical length.[16] The LCTV also was located behind the Fourier lens to further reduce the size of the JTC. In order to obtain phase-mostly modulation, the bias voltage was set to the lowest setting (0.978 V). The input scene was then binarized into two particular gray levels to produce a zero and π phase difference so that maximum modulation index on the JTPS was obtained. The recorded spectrum was not binarized, but the gray level of the fringes in the spectrum was reassigned to modulate between 0° phase and 210° phase to maximize the diffraction efficiency.* We stress that the gray scale modification was operated without the expense of additional computation time by using the look-up tables (the tables that map the input intensity to a gray scale value) built into the image frame grabber. We also note that the gray scale was modified to obtain maximum performance under a phase-only modulation assumption, but it may not be the optimum scheme if the polarization modulation effect is being considered as well.

7. DEMONSTRATION

To demonstrate the ability of the tracking algorithm that we have proposed, a motion sequence of three targets was generated, as shown in Fig. 3. A target tracking model can then be set up as follows: First, the dynamic state of the target at the k frame is defined as

$$x(k) = \begin{pmatrix} x(k) \\ \bar{x}(k) \end{pmatrix} . \qquad (13)$$

Since motion along each coordinate can be decoupled, only one direction needs to be considered. Acceleration and maneuvering of targets can be handled by including processing noise n in the dynamic equation:

$$x(k+1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x(k) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} n , \qquad (14)$$

and the measurement is

$$z(k) = (0 \quad 1)x(k) + w , \qquad (15)$$

where w is the measurement noise. A Kalman filtering process can therefore be applied to the tracking problem. The innovation covariance S can be computed, so the normalized squared distance D in Eq. (7) can be evaluated and the data association process described in Sec. 3 can be carried out. However, our goal here is to demonstrate the capability of the optical tracker and to illustrate the data association algorithm. Therefore, in

Fig. 2. A compact joint transform correlator architecture. The LCTV has no polarizer/analyzer attached. L1: collimating lens; L2: Fourier transform lens; L3: negative lens for beam compression. The linearly polarized light was rotated to align with the liquid crystal molecule director of the LCTV by using a halfwave plate.

*For a sinusoidal phase grating exp($i\phi\cos wx$) ($|\phi|$ = maximum modulation angle), the diffraction amplitude in the first order is proportional to the Bessel function $J_1(\phi)$, where $J_1(\phi)$ is maximum at $\phi \approx 105°$. By considering the fringe modulation in a JTPS as a sinusoidal grating, maximum diffraction efficiency is obtained by assigning the fringes to modulate between 0° and 210°

Fig. 3. Motion sequence of three targets: helicopter and jet, moving from bottom to top; airplane, stationary.
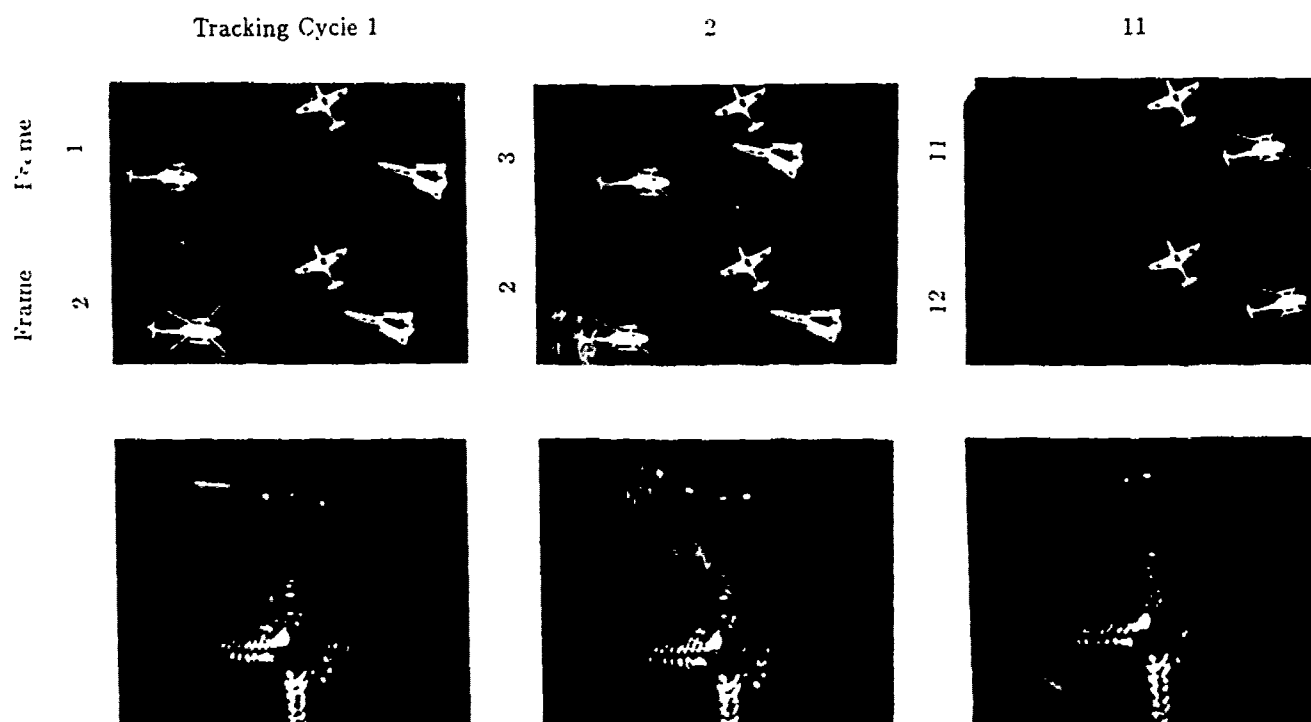
Tracking Cycle 1   2   11

Fig 4. (a) The image sequence as displayed on the LCTV at each step of the tracking cycle.
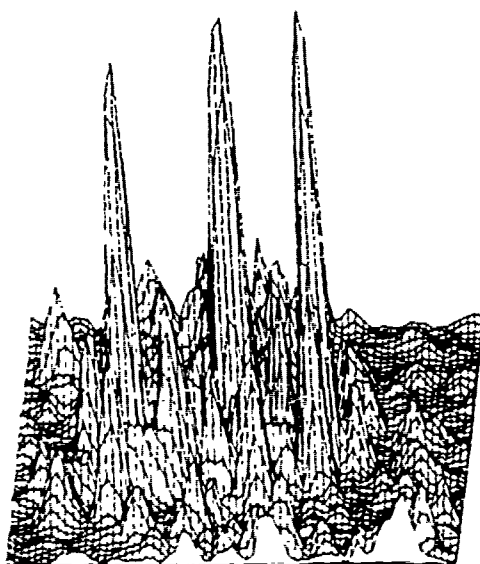(b) The correlation results as captured by the CCD camera at the corresponding tracking cycle.

Fig. 5. A 3-D plot of the correlation result obtained in one of the tracking cycles. Height of peaks are linear in intensity.

order to show the basic performance rather than the filtered results in this demonstration, we simply assume a noise-free measurement situation ($w = 0$). This immediately sets the gain of the filter to one. The update dynamic states of the target are then solely determined by the new measurements, and the predicted measurement is given by

$$ \hat{x}(k-1) = z(k-1) \tag{16} $$

For simplicity, the value of S is also assigned as the variance of the target's velocity in a number of tracking cycles.

A computer program was written that controls the joint transform process as well as the data association process based on the above simplification and the discussion in Sec. 3. To initiate the coordinate the system was first set to run with a stationary object as input, and the position of the correlation peak thus generated was assigned as the origin of the velocity plane coordinate. The tracking program was then set to run autonomously with the motion sequence in Fig. 4 as inputs. The computer program was written in Microsoft C, which ran on a microcomputer using an INTEL 386 processor operated at 25 MHz, and the processing time for a tracking cycle was approximately 0.7 s. Figures 4(a) and 4(b) depict the images as displayed on the LCTV and the corresponding correlation spots generated during each step of the tracking cycle. A 3-D plot of correlation peaks is also given, in Fig. 5. To avoid unnecessary data transfer, the current frame was alternately loaded onto the upper/lower half of the LCTV at every tracking cycle. This modification generates a correlation plane whose coordinate axes are inverted at each cycle, and the location of the measurements have to be interpreted accordingly. Table 1 summarizes the empirical data, and the final results are shown in Fig. 6, in which the actual locations of the targets at each step of the tracking cycle are also provided for comparison. It can be seen that a relatively simple hybrid optical architecture is capable of tracking multiple targets rather accurately.

8. CONCLUSION

We have demonstrated a data association multitarget tracking technique using a hybrid JTC. The role of the correlator is to

TABLE I. Empirical data after 11 tracking cycles.

	Tracking Cycle	1	2	3	4	5	6	7	8	9	10	11
a	Intensity[2]	190	188	183	188	187	121	183	117	187	174	188
	Location[3]	(-24,-1)	(32,16)	(1,0)	(32,13)	(1,-1)	(19,22)	(0,0)	(31,13)	(0,1)	(-21,8)	(2,-1)
	Path Assig[4]	I	III	II	III	II	ione	II	III	II	I	II
b	Intensity	183	190	185	185	189	186	152	189	178	187	120
	Location	(2,0)	(22,4)	(-23,0)	(23,3)	(22,5)	(30,10)	(20,7)	(-23,6)	(-20,7)	(0,0)	(-21,9)
	Path Assig.	II	I	I	I	I	III	I	I	I	II	I
c	Intensity	187	188	185	188	189	188	188	189	151	None	None
	Location	(32,13)	(-1,1)	(31,12)	(-1,1)	(30,14)	(-23,6)	(29,10)	(0,1)	(30,12)		
	Path Assig	III	II	III	II	III	I	III	II	III		
d	Intensity	None	None	None	None	None	186	None	None	114	None	None
	Location						(0,-1)			(-30,11)		
	Path Assig						II			None		

Measurement threshold level was set at 100 in 0-255 gray scale **Peak intensity*

Location of the peak was mesured in the velocity plane coordinate

Path assignments were made on-line based on the normalized square distance between the measurement in the previous cycle and that in the current cycle. Identifications between the paths and the corresponding targets were made through manual interaction after the experiment. Path I helicopter. II stationary aeroplane. III higher jet
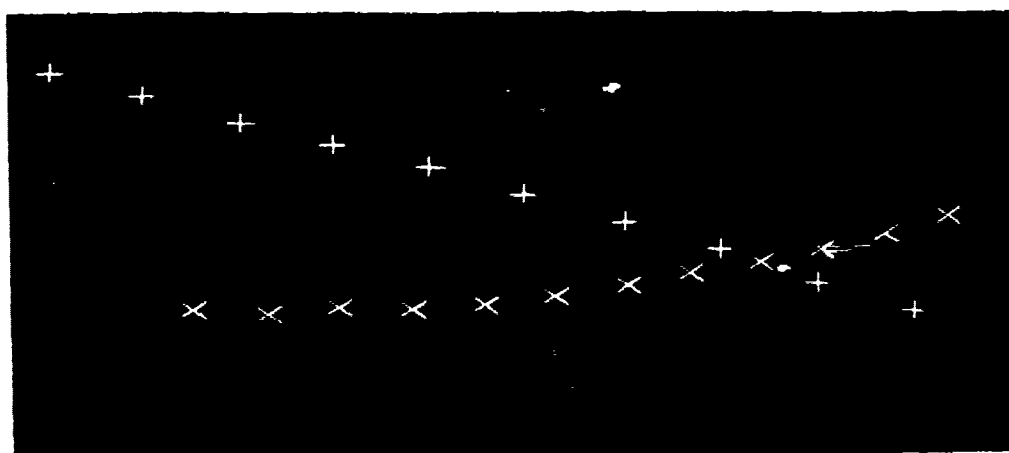
Fig. 6. Final results obtained after 11 tracking cycles. The actual locations of the targets at each of the tracking cycles are also given for comparison. The initial locations of the paths were assigned at the correct positions in this figure through manual interaction.

reduce the information of an input scene with large space-bandwidth product into a few correlation signals (velocity measurements) so that a digital system is capable of tracking multiple targets at real-time. Simplicity in optical architecture. relaxation of alignment requirements. and the adaptive property are the major advantages of a hybrid JTC in this mode of operation. The phase modulation property of a LCTV has been investigated and was employed in this experiment to improve the detection efficiency of a JTC. A motion sequence of three targets was used as the input scenes to demonstrate the applicability of this technique. and satisfactory results were obtained.

10. REFERENCES

1 A. VanderLugt. "Coherent optical processing. Proc. IEEE 64. 1300 (1974)
2 A. D. Gara. "Real time tracking of moving objects by optical correlation. Appl. Opt. 18, 172 (1979)

3. T. H. Chao and H. K. Liu, "Real-time optical holographic tracking of multiple objects," Appl. Opt. 28, 226 (1989).
4. C. S. Weaver and J. W. Goodman, "A technique for optically convolving functions," Appl. Opt. 5, 1248 (1966).
5. J. E. Rau, "Detection of differences in real distribution," JOSA 56, 1490 (1966).
6. E. C. Tam. F. T. S. Yu, D. A. Gregory, and R. Juday, "Autonomous real-time object tracking with an adaptive joint transform correlator," Opt Eng. 29(4), 314–320 (1990).
7. See, for example, P. Mayback. Stochastic Models, Estimation and Controls, Academic, New York (1979).
8. F. T. S. Yu, Q. W. Song, Y. S. Cheng, and D. A. Gregory, "Comparison of detection efficiencies for VanderLugt and joint transform correlators," Appl. Opt. 29, 225–232 (1990).
9. F. T. S. Yu and Q. W. Song, "Hard-clipping joint transform correlator using a microchannel spatial light modulator," Micro. Opt. Tech. 1, 323–326 (1988).
10. D. Psaltis, E. G. Paek, and S. S. Venkatesh, "Optical image correlation with a binary spatial light modulator," Opt. Eng. 23(6), 698–704 (1984).
11. J. A. Davis, R. A. Lilly, K. D. Krenz, and H.-K. Liu, "Applicability of the liquid crystal television for optical data processing," in Nonlinear Optics and Applications, Proc. SPIE 613, 245–254 (1986).
12. L. M. Blinov, Electro-Optical and Magneto-Optical Properties of Liquid Crystals, Chap. 4, Wiley, New York (1983).
13. N. Konforti, E. Marom, and S. T. Wu, "Phase-only modulation with twisted nematic liquid crystal spatial light modulator," Opt. Lett. 13, 251–253 (1988).
14. T. H. Barnes, T. Eiju, K. Matsuda, and N. Ooyama, "Phase-only modulation using a twisted nematic liquid crystal television," Appl. Opt. 28, 4845–4852 (1989).
15. D. A. Gregory, J. A. Loudin, J. Kirsch, E. C. Tam, and F. T. S. Yu, "Application of hybrid modulation properties of the liquid crystal television," submitted to Appl. Opt.
16. D. L. Flannery, A. M. Biernacki, J. S. Loomis, and S. L. Cartwright, "Real-time coherent correlator using binary magneto-optic spatial light modulators at input and Fourier planes," Appl. Opt. 25, 466 (1986).

Eddy C. Tam received his BS degree in physics in 1984 from the State University of New York College at Buffalo and his MS degree in physics in 1987 and his Ph.D degree in electrical engineering in 1990 from the Pennsylvania State University. His research interests include optical image processing and pattern recognition, holography, and spatial light modulators. He is currently an assistant professor in physics at the University of North Carolina at Charlotte.

Francis T. S. Yu: Biography and photograph appear with the paper "Self-organizing optical neural network for unsupervised learning" in this issue.

Aris Tanone received his BS degree and doctor candidate degrees in physics from Gadjah Mada University in 1974 and 1976, respectively. He got his MS degree in applied physics engineering from Hokkaido University in 1981. In 1982 he joined the graduate school of opto-electrotechniques and laser application at the University of Indonesia as a research associate, and then he spent around three years in the microcomputer industry. In 1986 he became the Vice Dean of Academic Affairs at the College of Engineering, Satya Wacana Christian University in Indonesia. He came to the Pennsylvania State University to pursue his Ph.D. degree in electrical engineering in 1989. His current research interests include optical image processing, pattern recognition, holography, and spatial light modulators. He is a member of OSA SPIE, Indonesian Physics Society, and Japan Society of Applied Physics.

Don A. Gregory: Biography and photograph appear with the paper "Self-organizing optical neural network for unsupervised learning" in this issue.

Richard D. Juday holds a BA in physics and a Ph.D. in electrical engineering. Among other areas he has worked in radiation environments for manned space flight and earth resources remote sensing. He currently works in digital and optical image processing for machine vision. Video-rate image coordinate transformations, one aspect of his machine vision activities, are applied to human low vision problems. He has avid interests in whitewater boating, badminton, and Mexican mountain-stomping and canyon-crawling.

APPENDIX 10.18

Unsupervised Neural Net

Self-organizing optical neural network for unsupervised learning

Taiwei Lu, MEMBER SPIE
Francis T. S. Yu, FELLOW SPIE
The Pennsylvania State University
Electrical and Computer Engineering
    Department
University Park, Pennsylvania 16802

Don A. Gregory
U.S. Army Missile Command
Redstone Arsenal, Alabama 35898

Abstract. One of the features in neural computing must be the ability to adapt to a changeable environment and to recognize unknown objects. This paper deals with an adaptive optical neural network using Kohonen's self-organizing feature map algorithm for unsupervised learning. A compact optical neural network of 64 neurons using liquid crystal televisions is used for this study. To test the performance of the self-organizing neural network, experimental demonstrations and computer simulations are provided. Effects due to unsupervised learning parameters are analyzed. We show that the optical neural network is capable of performing both unsupervised learning and pattern recognition operations simultaneously, by setting two matching scores in the learning algorithm. By using a slower learning rate, the construction of the memory matrix becomes more organized topologically. Moreover, the introduction of forbidden regions in the memory space enables the neural network to learn new patterns without erasing the old ones.

CONTENTS

1. INTRODUCTION

It is well known that digital computers can solve computational problems millions of times faster than human brains. However, for cognitive tasks, such as recognition, learning, and adaptation in a complex environment, the performance of the human brain is beyond the reach of electronic computers. The human brain consists of millions of neurons, with each neuron interconnected to thousands of other neurons. Artificial neural networks (ANNs) have adopted this unique feature of massive interconnections, simulating the learning capability of a biological system.[1-9]

Strictly speaking, two types of learning processes are used in the human brain: supervised and unsupervised learning.[7] If a teacher first organizes the information and then teaches the students, the learning process is obviously supervised. For example, in an artificial neural network, both the input and desired output data must be provided as training examples. In other words, the ANN has to be taught when to learn and when to process the information. Nevertheless, if an unknown object is presented to the artificial neural network during the processing, the network may provide an erroneous output result.

On the other hand, for unsupervised learning (also called self-learning), the students learn by themselves, relying on some simple rules and their past experiences. In an ANN, only the input data are provided, while the desired output result is not. After a single trial or series of trials, an evaluation rule (previously provided to the neural network), is used to evaluate the performance of the network. Thus, we see that the network is capable of adapting and categorizing unknown objects. This kind of self-organizing process is a representation of the self-learning ability of the human brain.

In past decades, models of ANNs have been developed. To name a few, the perceptron, the error-driven backpropagation,[3] the Hopfield models, and the Boltzmann machine are examples of the supervised learning models. The adaptive resonance theory (ART),[4] the neocognitron,[5] the madline, and the Kohonen self-organizing feature map[6] are among the best known unsupervised learning models. Among these models, the Kohonen model is the simplest self-organizing algorithm, capable of performing statistical pattern recognition and classification, and it can be modified for optical neural network implementation. However, Kohonen's model was originally designed for feature mapping of statistical data: to perform adaptive pattern recognition, we define three matching parameters and incorporate matching criteria in the learning algorithm. The effects of fast and slow learning are analyzed. The concept of forbidden regions in the memory space is introduced in the Kohonen model such that the neural network learns new input objects without erasing old memory.

The development of ANNs has recently stimulated great interest for optical implementations.[10-16] Because of large information capacity and the capability of establishing three-dimensional massive interconnections, optics is considered a potential candidate for constructing massively parallel neural networks. The first optical neural network, based on a vector-matrix processor,

Output space: MxM output neurons

Memory space:

Fig. 3. Neighborhood selection in the memory matrix space using Kohonen's self-learning algorithm: the initial neighboring region is large. As time proceeds, the region shrinks until it reduces to one memory sub-matrix.
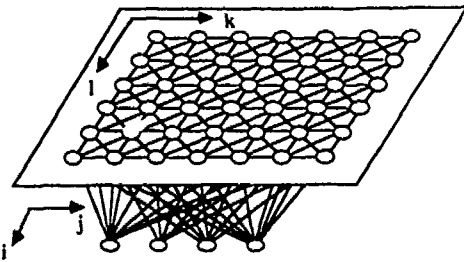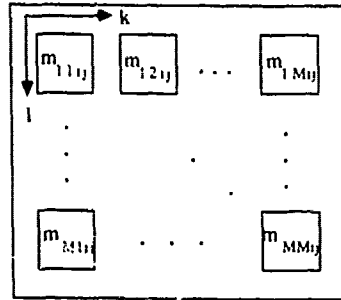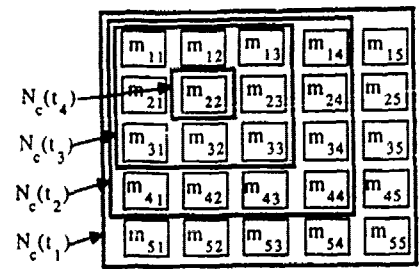
Input space: NxN input neurons

Fig. 1 Single layer neural network.

Fig. 2. Memory vectors in the memory matrix space.

was proposed by Farhat and Psaltis and their coworkers in 1985.[10] The optical neural network is composed of a lenslet array to establish the optical interconnections between the input light emitting diode (LED) array and the interconnection weight matrix (IWM). However, the architecture required a high resolution spatial light modulator (SLM), which is expensive to build and may not be available in the near future. To alleviate these shortcomings, we developed an optical neural network using an inexpensive video monitor for displaying a programmable IWM.[14,15] Recently, a compact optical neural network using pocket size liquid crystal televisions (LCTVs) has also been built.[16] We shall use Kohonen's self-organizing algorithm for the unsupervised learning in one compact optical architecture.

2. KOHONEN'S SELF-ORGANIZING FEATURE MAP ALGORITHM

Knowledge representation in human brains is generally at different levels of abstraction and assumes the form of a feature map in the real world. The Kohonen model suggests a simple learning rule by adjusting the interconnection weights between input and output neurons based on the matching score between the input and the memory. A single layer neural network is defined that consists of $N \times N$ input and $M \times M$ output neurons forming an input and output vector space, as illustrated in Fig. 1. Notice that the output neurons are extensively interconnected to the input neurons with weighted interconnections. The output nodes are also laterally interconnected with each other. Let us assume that 2-D vectors (i.e., input patterns) are sequentially presented to the neural network, given by

$$x(t) = \{x_{ij}(t), \quad i,j = 1,2,....,N\},$$ (1)

where t represents the time index, such as the iteration number in discrete-time sequence, and (i,j) specifies the position of the input neuron. Thus, the output vectors of the neural network can be expressed as the weighted sum of the input vectors

$$y_{lk}(t) = \sum_{i=1}^{N}\sum_{j=1}^{N} m_{lkij}(t)x_{ij}(t), \quad l,k = 1,2,....,M,$$ (2)

where $y_{lk}(t)$ represents the state of the (l,k)th neuron in the output space and $m_{lkij}(t)$ is the interconnection weight between the (ij)th input and the (l,k)th output neurons. Equation (2) can also be written in matrix inner-product representation given by

$$y_{lk}(t) = m_{lk}(t)x(t),$$ (3)

where $m_{lk}(t)$ can be considered to be a two-dimensional vector in a 4-D memory matrix space, which can be expanded in an array of 2-D submatrices, as shown in Fig. 2. Each submatrix can be written in the form

$$m_{lk}(t) = \begin{bmatrix} m_{lk11}(t) & m_{lk12}(t) & \cdots & m_{lk1N}(t) \\ m_{lk21}(t) & m_{lk22}(t) & \cdots & m_{lk2N}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ m_{lkN1}(t) & m_{lkN2}(t) & \cdots & m_{lkNN}(t) \end{bmatrix},$$ (4)

where the elements in each submatrix represent the associative weight factors from each of the input neurons to one output neuron.

Notice that the Kohonen model does not in general specify the desired output results. Instead, a similarity matching criterion is defined to find the best match between the input vector and the memory vectors and then determine the best matching output node. The optimum matching score d_c as defined by Kohonen can be written as

$$d_c(t) = \|x(t) - m_c(t)\|$$ (5)

$$= \min_{l,k}\{d_{lk}(t)\} = \min_{l,k}\{\|x(t) - m_{lk}(t)\|\},$$ (6)

where $c = (l,k)^*$ represents the node in the output vector space at which a best match occurred. $\|\cdot\|$ denotes the Euclidean distance operator.

After obtaining an optimum matching position c, a neighborhood $N_c(t)$ around the node c is further defined for IWM modification, as shown in Fig. 3. Notice that the memory matrix space is equivalent to the output space, in which each submatrix corresponds to an output node. As time progresses, the neighborhood $N_c(t)$ will slowly reduce to the neighborhood that consists of only the selected memory vector m_c, as illustrated in the figure.

Furthermore, a simple algorithm is used to update the weight factors in the neighborhood topology, by which the similarity between the stored memory matrix $m_{lk}(t)$ and the input vector $x(t)$ increases. Notice that the input vectors can be either binary or analog patterns, while the memory vectors are updated in an analog incremental process. The adaptation formula of the algorithm can be written as
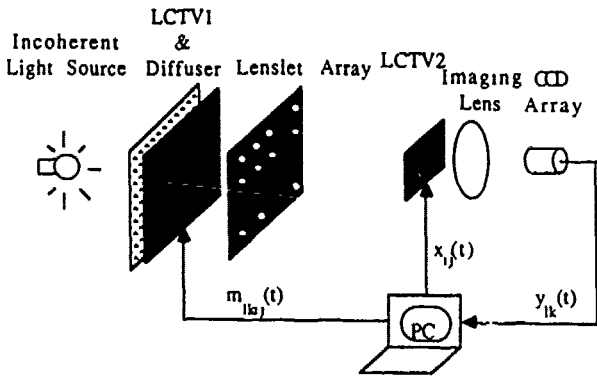
Fig. 4. Schematic diagram of a self-organizing optical neural network.

Fig. 5. Experimental setup of a self-organizing optical neural network.

$$m_{lk}(t+1) = \begin{cases} m_{lk}(t) + a(t)[x(t) - m_{lk}(t)] , & \text{for } l,k \in N_c(t) , \\ m_{lk}(t) , & \text{otherwise} . \end{cases} \quad (7)$$

where $0 < a(t) < 1$ represents a gain sequence, called the learning speed, that is usually a monotonically slowly decreasing function of t. In practice, the learning speed is assumed to be linear, which can be written as

$$a(t) = a(0) - \alpha t , \quad (8)$$

where α is a learning rate.

Mention must be made that the point density of the memory vectors tends to approximate the probability density function of the input vectors, which has been proved by Kohonen.[6]

It should be noted that Eq. (6) can be expanded in the following form:

$$d_{lk} = \sum_{i=1}^{N} \sum_{j=1}^{N} [x_{ij}(t) - m_{lkij}(t)]^2$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij}^2(t) + \sum_{i=1}^{N} \sum_{j=1}^{N} m_{lkij}^2(t) - 2 \sum_{i=1}^{N} \sum_{j=1}^{N} m_{lkij}(t)x_{ij}(t) , \quad (9)$$

where the first term is a constant with regard to the output position (l,k). If the weight vectors are normalized so that their autocorrelations (i.e., the sum of the squared weights from all inputs to each output node) are identical, then the second term also becomes a constant. Under this condition, the minimum Euclidean distance occurs whenever the third term becomes maximum, i.e.,

$$d_c(t) = \min_{l k}\{d_{lk}(t)\} = \max_{l k}\left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} m_{lkij}(t)x_{ij}(t) \right\} , \quad (10)$$

which represents the output result of the neural network. In other words, selecting the minimum Euclidean distance between the memory vectors and the input pattern vector is equivalent to finding the maximum output node in the output space.

The maximum output node can be determined by using extensive lateral inhibition among the output nodes in the output space, which is known as the MAXNET algorithm.[2] The expression of the inhibit interconnections can be written as

$$y_{lk}(t+1) = f\left[y_{lk}(t) - \varepsilon \sum_{\substack{m=1 \\ (m,n) \neq (l,k)}}^{M} \sum_{n=1}^{M} y_{mn}(t) \right] , \quad l,k = 1,2, ...,M . \quad (11)$$

where f is a nonlinear operator over the output nodes and ε is an inhibition constant between the output nodes. For simplicity, we assume that the nonlinear operation $f(x)$ is a thresholding function, given by

$$f(x) = \begin{cases} 1 , & x \geq a , \\ x , & b \leq x < a , \\ -1 , & x < b , \end{cases} \quad (12)$$

where a and b are arbitrary constants. By referring to the MAX-NET algorithm, we see that each node will excite itself and inhibit the other nodes during each iteration. Finally, only the maximum output node will retain, while all of the other nodes will eventually go to zero. Since the MAXNET algorithm will always converge, for $\varepsilon < 1/M$, the maximum node can always be found.[2]

3. LCTV OPTICAL NEURAL NETWORK

Recent developments in liquid crystal display technology have made the imaging quality approach that of commercially available video monitors. For instance, the contrast ratio of a recent LCTV, with built-in thin-film transistors (TFTs), is measured at about 15:1 under white light illumination (it can be higher than 30:1 if illuminated by laser light[17]); the dynamic range is about 16 gray levels, which can be continuously adjusted; and the transmittance of the LCTVs is in the range of 5% to 10%. Thus, the LCTVs can be used to build inexpensive optical neural networks.

The schematic diagram of an LCTV optical neural network and the experimental setup are shown in Figs. 4 and 5. A neural network of $8 \times 8 = 64$ interconnected neurons was recently constructed[16] for which an 80 W xenon arc lamp was used as the incoherent light source. In this architecture, a Hitachi LCTV (LCTV1) is used for the generation of the IWM, which consists of an 8×8 array of submatrices, and each submatrix has 8×8 elements. This IWM is then displayed on a fine diffuser immediately behind LCTV1. A Seiko LCTV (LCTV2) is used as an input device for the generation of the input patterns. The
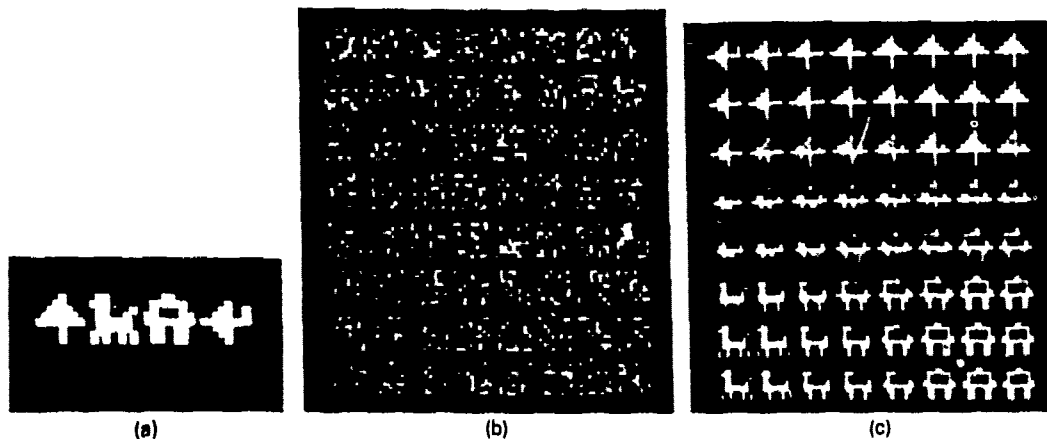
Fig. 6. Experimental demonstration of a self-organizing optical neural network using unsupervised learning: (a) Input patterns. (b) Initial memory matrix space with random noise. (c) Final memory matrix space. The four patterns are adapted in the memory matrix space; the tree is centered at (1,8), the dog at (7,1), the house at (7,7), and the airplane at (1,1).

lenslet array, which consists of 8×8 lenses, provides the interconnections between the IWM and the input pattern. Each lens of the lenslet array images each of the IWM submatrices onto the input LCTV2 to establish the proper interconnections. Thus, the IWM submatrices can be superimposed to illuminate the input vector pattern of LCTV2. This represents the matrix-vector product operation [i.e., the interconnective part of Eq. (10)]. The output result from LCTV2 is then picked up by a CCD array detector for MAXNET operation. The final result can be obtained by a microcomputer. Notice that the data flow in the optical system is controlled primarily by the computer. For instance, the IWMs and the input patterns can be written onto LCTV1 and LCTV2, and the computer can also make decisions based on the output results of the neural network. Thus, the proposed LCTV optical neural network is a programmable and adaptive neural network.

4. UNSUPERVISF⁰ LEARNING IN THE OPTICAL NEURAL NET⁰⁰ K

Before apply·   ·-lf-organizing model in the optical neural network, tw· ·   ·⸰actors must be considered: one is the system components ·.·. al. nent errors, the other is the effect of the parameters of the sei. organizing system.

Since the IWM submatrices are precisely interconnected with the input pattern vector by a lenslet array, the alignment of the interconnections is rather critical. Although the adjustment of lenses is formidable, it is rather simple to shift the IWM submatrices in LCTV1. A set of test patterns can be displayed on LCTV1 and LCTV2. The test patterns are shifted in small steps according to the sharpness of the output pattern detected by the CCD camera. Thus, the interconnection alignment can be self-adjusted via the feedback loop.

The uniformity and stability of the light illumination also pose some problems on the accuracy of the output results. To alleviate the uniformity problem, a test is performed when a new input pattern vector is presented. During the test, all elements in the memory matrix are set to the same value, i.e.,

$$m_{lkij}(t_0) = 1 , \quad l,k, = 1,2,...,M , \quad i,j, = 1,2,...,N . \quad (13)$$

Then an output array can be obtained, such as

$$\hat{y}_{lk}(t_0) = \sum_{i=1}^{N}\sum_{j=1}^{N} |x_{ij}(t_0)| , \quad l,k = 1,2,...,M . \quad (14)$$

The uniformity test array $\hat{Y}(t_0)$ is then divided by the output array during the self-learning and recognition processes, given by

$$\overset{*}{y}_{lk}(t) = \frac{y_{lk}(t)}{\hat{y}_{lk}(t_0)} , \quad l,k = 1,2,...,M . \quad (15)$$

where $Y^*(t)$ is the normalized output array in which the non-uniformity for the light illumination can be eliminated.

Another problem with the light source is the fluctuation of the arc lamp, which causes detection errors. To overcome this problem, we took a sequence of output frames and performed average operations for every iteration.

For experimental demonstrations, four 8×8 pixel binary input patterns (i.e., a tree, a dog, a house, and an airplane) were sequentially presented at the input LCTV2 shown in Fig. 6(a). Figure 6(b) shows the initial memory matrix (i.e., IWM) as a random pattern. The first neighboring region is chosen as $N_c(0) = 5$, the initial learning speed is $a(0) = 0.01$, and the learning rate is selected as in computer simulations, which is $\alpha = 0.000025$ per iteration. The output pattern picked up by the CCD camera must be normalized by referring to Eq. (15). Then the location of the maximum output intensity can be identified by using the MAXNET algorithm. The memory submatrices in the neighborhood of the maximum output spot are adjusted based on the adaptation rule of the Kohonon model. The amount of analog increase in the memory vectors can be controlled by the learning rate α. The updated memory matrix is then displayed on LCTV1 for the next iteration, and so on. Due to the gray level limitation in LCTV displays, a few iterations in the learning process may result in one noticeable gray level change by the CCD detector array. Since the inputs are binary patterns at this stage, the center memory vector in the neighborhood region eventually converges to the binary input vector. Thus, the limited dynamic range of LCTVs would not pose a major problem in our experiment. As
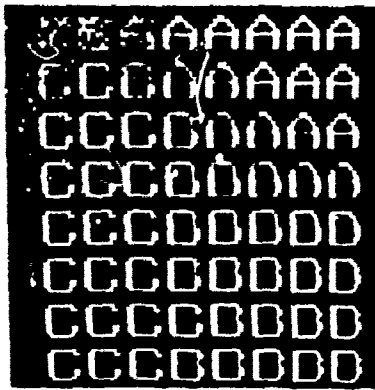
Fig. 7. Feature map in memory matrix space. Three similar patterns are located on the same side in the memory matrix, B centered at (8,4), C at (6,1), and D at (8,8), while pattern A sits in the opposite direction of the memory matrix, centered at (1,8).

Fig. 8. Fast and slow linear learning speeds.

shown in Fig. 6(c), after 400 iterations the memory has learned the four input patterns, mapping them around four overlapping regions in the IWM. The centers of these patterns are located at (1,8), (7,1), (7,7), and (1,1), respectively.

One of the major advantages in adaptive pattern recognition is to learn the unknowns in addition to recognizing known objects. This task can be achieved by adding three criteria to the Kohonen learning rule, for which the matching rate R for pattern recognition can be defined as

$$R = \frac{\|Y_c(t)\|}{\|X(t)\|} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} m_{cij}(t)x_{ij}(t)}{\sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}^2(t)} , \quad (16)$$

where $\|\cdot\|$ is the norm operator. Let us assume that two constants F_1 and F_2 represent the matching and unmatching scores, by which $F_1 < 1$, $F_2 < 1$, and $F_1 > F_2$; then the criteria of the learning rules are as follows: (a) If $R \geq F_1$, the input object matches the memory submatrix at position c. Then the neural network recognizes the input object. (b) If $F_2 \leq R < F_1$, the input object is identified as a pattern that belongs to the class of patterns (i.e., submatrices) around the location c. This shows that the input object has some differences from the existing patterns in the class. Then the memory submatrices in the neighborhood of c should be modified to adapt to the new input pattern. (c) If $R < F_2$, the input object is considered to be an unknown pattern. Then a new category is developed to learn the unknown pattern using Eq. (7).

For illustration, F_1 and F_2 are selected as $F_1 = 95\%$ and $F_2 = 75\%$. A computer-simulated self-learning process is illustrated in Fig. 7, in which four standardized English letters A, B, C, and D are sequentially presented as the input objects to the neural network. It can be seen that in the memory matrix, the locations of the submatrices reflect the similarity among input patterns. The grouping of the submatrix locations is known as Kohonen's feature map, for which the patterns of similar features tend to stay close to each other. For example, B, C, and D are similar patterns; their submatrices are located on the same side of the memory matrix, centered at (8,4), (6,1), and (8,8), respectively, while pattern A sits in the opposite direction, located at (1,8).
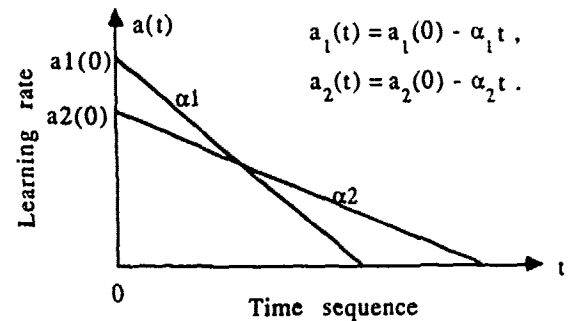
It should be noted that the innerproducts of the 4-D memory matrix and the 2-D input pattern are performed in parallel by the optics, while the MAXNET algorithm, the thresholding, and the adaptive operations are performed by a microcomputer.

Our experiences have shown that the best results in self-organization are obtained when the neighborhood of the memory matrix space is first selected as a wide region and then allowed to shrink as time proceeds. The linearly decreasing learning speed $a(t)$ is illustrated in Fig. 8. Since $\alpha_1 > \alpha_2$, it is clear that $a_1(t)$ is a faster learning speed than $a_2(t)$. It is apparent that to reach the stable state of the memory matrix, $a_1(t)$ would require fewer iterations than $a_2(t)$. However, in real life, the intensive training may not always be effective. With the slower learning speed, the memory matrix is more organized and also appears smoother. The price paid is obviously the training time required for using the slower learning speed. As an example, the memory matrix space adapted from the faster learning speed is shown in Fig. 9(a). After 400 iterations, using $a(0) = 0.02$ and $\alpha = 0.00005$, the similar patterns I, J, and T are adapted in the lower left part of the memory matrix, centered at (5,1), (7,2), and (8,4), respectively. The pattern X submatrices occupy the upper right part, centered at (3,8). Since pattern X is very different from the I, J, and T patterns in the memory space, using slower learning rate, $\alpha = 0.000025$, pushes the pattern X submatrices farther away to the upper right corner, centered at (1,8) after 800 iterations, as shown in Fig. 9(b). In this figure, we see that the memory space becomes more organized topologically.

When the similar patterns are presented at the input of a neural network, a new pattern may occasionally override the old pattern through selecting the same optimum position in the memory matrix space, with the old pattern gradually fading away. This phenomenon is rather similar to the learning process of humans: if one moves to Southern California for a long period of time, he may forget the bitterness of the cold winters in the northeast region of the United States.

An example of this phenomenon has been simulated on computer, as shown in Fig. 10. An input pattern B is first presented to the optical neural network. The memory matrix adapts the pattern around the lower right corner, centered around (8,8), as shown in Fig. 10(a). A second pattern E is presented to the neural network at a later time. Since pattern E is similar to pattern B, it picks the same spot (8,8) in the memory matrix space as its maximum output point. As can be seen in Fig. 10(b), the memory submatrices of B are eventually taken over by E, here after 100 iterations. If one would like to preserve the old memory while learning new knowledge, there are two ways to
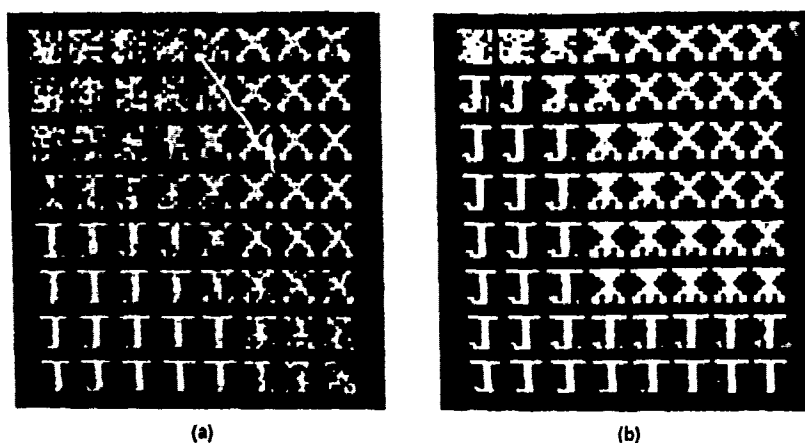
Fig. 9. Memory matrix topology using linear learning speeds. (a) Fast learning rate with the centers of the similar patterns I at (5,1), J at (7,2), and T at (8,4) and the center of the different pattern X at (3,8). (b) Slow learning rate with the pattern X pushed farther away to the upper right corner, centered at (1,8). Notice that the memory space becomes topologically more organized and also appears smoother.

achieve this. One way is to refresh the memory by repeatedly reviewing the previous input patterns, which takes a lot of computational time and may cause confusion. Another is to set a rule in the learning process: the new pattern would not take the output nodes within certain regions of the old patterns in the memory matrix, i.e., their center spots. This rule can be expressed as follows:

$$d_c(t_n) = \max_{(l,k) \in R_e(t_n)} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} m_{lkij}(t_n) x_{ij}(t_n) \right\} , \qquad (17)$$

where $R_e(t_n)$ consists of the output nodes of $d_c(t_1)$, $d_c(t_2)$, ..., $d_c(t_{n-1})$ and their neighboring regions, illustrated in Fig. 10(c). The radii r_e of the forbidden region can be determined by experience (in our experiments, $r_e = 2$ was selected). After this rule is added to the Kohonen model, the neural network can easily remember 11 patterns without erasing similar patterns, as shown in Fig. 10(d). The centers of these 11 patterns in the memory matrix are B at (8,4), C at (8,6), E at (6,8), F at (8,8), I at (1,8), J at (1,6), T at (3,8), V at (7,1), W at (6,3), X at (1,1), and Y at (1,3).

Although results can be obtained in the self-organizing optical neural network by adjusting the learning parameters, the best results were achieved through experience. In addition, since the original Kohonen model was not designed for pattern recognition, the ability for self-organization and adaptive pattern recognition can be enhanced by incorporation with more effective models, such as the interpattern association (IPA) model,[18] in constructing the interconnection weight matrix during the learning process.

5. CONCLUDING REMARKS

We have demonstrated that the Kohonen self-organizing feature map algorithm can be implemented in an optical neural network to perform adaptive pattern recognition. This model is modified and combined with several matching criteria we have developed. Experimental and computer-simulated results have shown that

the self-organizing optical neural network has the ability to develop new categories in the memory matrix space to learn unknown objects while performing pattern recognition operations. The optical neural network is also capable of organizing a feature map and preserving old memory while learning new knowledge. These abilities can be achieved by setting new criteria of matching scores. using a slower learning speed, and introducing the concept of forbidden regions in the associative memory space.

## 6. ACKNOWLEDGMENT

7. REFERENCES

1. R. J. Baron, *Cerebral Computer. An Introduction to the Computational Structure of the Human Brain*. Lawrence Erlbaum Assoc. Publ.. Hillsdale, N.J. (1987)
2. R. P. Lippmann. "Introduction to computing with neural nets." IEEE ASSP Mag.. pp. 4–22 (Apr. 1987).
3. D E Rumelhart and J. L. McClelland. eds.. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vols. 1 & 2. MIT Press. Cambridge (1986)
4. G A Carpenter and S Grossberg. "Massively parallel architecture for a self-organizing neural pattern recognition machine." Comput. Vision Graph. Image Proc. 37. 54 (1987).
5. K Fukushima. "Neural network for visual pattern recognition." Comput. 21(3). 65–115 (1988).
6. K Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag. New York (1984).
7. V Vemuri. "Artificial neural networks an introduction." in *Neural Networks—Artificial Neural Networks. Theoretical Concepts*. V Vemuri. ed.. pp. 1–11. IEEE Comput. Soc. Press (1988).
8. J J Hopfield. "Artificial neural networks.' IEEE Circ. Devices Mag.. pp. 3–10 (Sept 1988).
9. R Hecht-Nielsen. "Neurocomputing. picking the human brain." IEEE Spectrum. pp. 36–41 (March 1988)
10. N. H. Farhat. D Psaltis. A. Prata. and E. Paek. "Optical implementation of the Hopfield model." Appl. Opt. 24(10). 1469 (1985).
11. N H Farhat. "Optoelectronic neural networks and learning machines." IEEE Circ Devices Mag.. pp 32–41 (Sept 1989).
12. Y S Abu-Mostafa and D Psaltis. "Optical neural computers." Sci Am.. pp. 88–95 (March 1987)
13. N H Farhat and D Psaltis. 'Optical implementation of associative memory based on models of neural networks." in *Optical Signal Processing*. J L. Horner. ed.. pp 129–162. Academic Press. New York (1987)
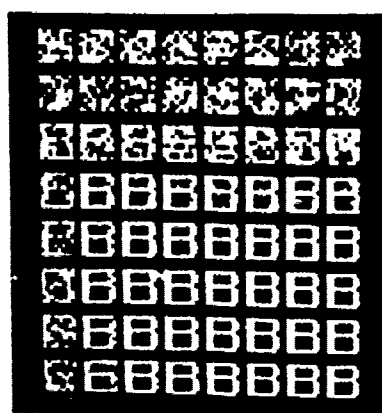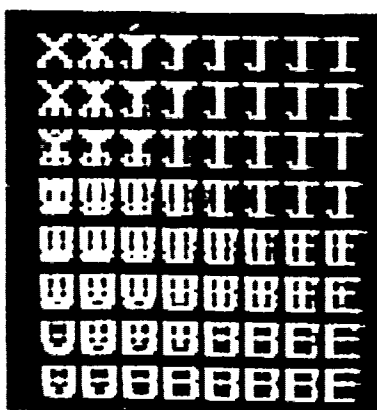
(a)
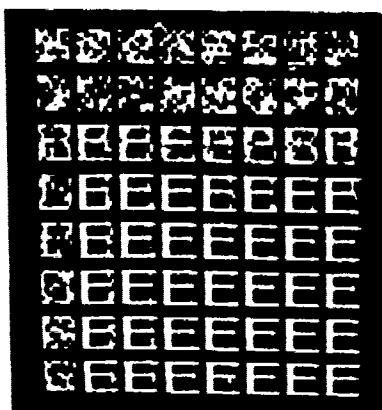
(b)
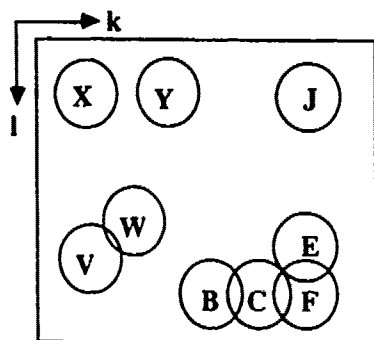
(c)

(d)

Fig. 10. Topological memory matrix space showing learning without erasing the old memory: (a) Pattern B in the memory matrix space. (b) Pattern B submatrices are erased, after pattern E is learned. (c) Forbidden regions in memory matrix space. (d) Learning new patterns without erasing old ones. The centers of the 11 patterns are B at (8,4), C at (8,6), E at (6,8), F at (8,8), I at (1,8), J at (1,6), T at (3,8), V at (7,1), W at (6,3), X at (1,1), and Y at (1,3).

14. S. Wu, T. Lu, X. Xu, and F. T. S. Yu, "Adaptive optical neural network using a high resolution video monitor," Microwave Opt. Technol. Lett. 2(7), 252 (1989).
15. T. Lu, S. Wu, X. Xu, and F. T. S. Yu, "Two-dimensional programmable optical neural network," Appl. Opt. 28(22), 4908 (1989)
16. F. T. S. Yu, T. Lu, X. Y. Yang, and D. A. Gregory, "Optical neural network using pocket size liquid crystal televisions," to be published in Opt. Lett.
17. H. K. Liu and T. H. Chao, "Liquid crystal television spatial light modulator," Appl. Opt. 28(22), 4722 (1989)
18. T. Lu, X. Xu, S. Wu, and F. T. S. Yu, "Neural network model using interpattern association," Appl. Opt. 29(2), 284 (1990).

Taiwei Lu received his BS degree from the University of Science and Technology of China in 1985 and his Ph.D. degree from the Pennsylvania State University in 1990, both in electrical engineering. He is currently a research scientist at Physical Optics Corporation, Torrance, Calif. His research interests include optical neural networks, optical interconnections, optoelectronic devices, machine vision, and digital/optical signal processing. He is a member of IEEE, OSA, SPIE, International Neural Network Society, and Tau Beta Pi.

Francis T. S. Yu received his BSEE from Mapua Institute of Technology, Manila, Philippines, and his MS and Ph.D. degrees in electrical engineering from the University of Michigan. From 1958 to 1965 he was a teaching fellow, instructor, and lecturer in the Electrical Engineering Department and a research associate with the Communication Sciences Laboratory at the University of Michigan. From 1965 to 1980 he was on the faculty of the Electrical and Computer Engineering Department at Wayne State University. He was a visiting professor in electrical and computer engineering at the University of Michigan during 1978–1979. Since 1980, Dr. Yu has been a professor in the Electrical Engineering Department at the Pennsylvania State University. He has been a consultant to several industrial and government laboratories. He is active in research in the fields of optical signal processing, holography, optics, and information theory and has published numerous technical papers in these areas. He was a recipient of the 1983 Faculty Scholar Medal for Outstanding Achievement in Physical Sciences and Engineering and the 1984 Outstanding Researcher in the College of Engineering at Penn State and was named Evan Pugh Professor of Electrical Engineering there in 1985. He is the author of *Introduction to Diffraction, Information Processing and Holography; Optics and Information Theory; Optical Information Processing; White-Light Optical Processing,* and *Principles of Optical Engineering.* Dr. Yu is a Fellow of OSA, SPIE, and IEEE.
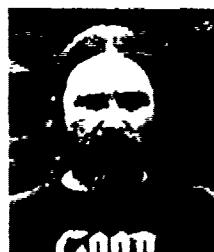
Don A. Gregory was born in Alabama in 1953. He earned BS and MS degrees in physics and a Ph.D. in physics in 1984 from the University of Alabama in Huntsville. Dr. Gregory is currently a research physicist at the U.S. Army Missile Command at Redstone Arsenal, Ala. His interests include optical computing, pattern recognition, and teaching. He is an adjunct professor of physics at the University of Alabama in Huntsville and the University of Alabama in Birmingham. He is also an adjunct professor of electrical engineering at the Pennsylvania State University.

APPENDIX 10.19

Color Holography Using LiNbO$_3$

Color holographic storage in LiNbO₃

Francis T.S. Yu, Shudong Wu, Andy Mayers and Sumati Rajan

The Pennsylvania State University, Department of Electrical Engineering, University Park, PA 16802, USA

and

Don A. Gregory

The U.S. Army Missile Command, Research Directorate, Redstone Arsenal, AL 35898-5248, USA

Experimental demonstrations of real-time color holographic storage in LiNbO₃ using a "white light" laser are presented. The geometry of the recording setup and wavelength crosstalk are discussed.

Two of the most widely used white-light holograms must be the reflection hologram of Denisyuk [1] and the rainbow hologram of Benton [2,3]. In reflection hologram, a thickness emulsion of about 20 μm would have a wavelength selectivity about $\Delta\lambda/\lambda = 1/40$, which is high enough to produce color hologram images without significant color blur. However, the physical requirements for constructing a reflection hologram is rather stringent, which prevents its wide spread use of applications. On the other hand, construction of a rainbow hologram requires a narrow slit, for which the parallax information of the hologram image would be partly lost.

In this communication, we shall demonstrate that color holograms can be constructed in a photorefractive crystal using a "white-light" laser. Since photorefractive crystal is much thicker than conventional photographic emulsion, it provides a higher wavelength selectivity such that the color blur can be minimized. Furthermore, the construction of photorefractive holograms is in real-time mode and the shrinkage of the emulsion can be prevented. As in contrast with the photographic film, multiplexing color holograms in a photorefractive crystal is possible.

By applying the coupled wave theory [4] in thick emulsion hologram, as illustrated in fig. 1, wave-
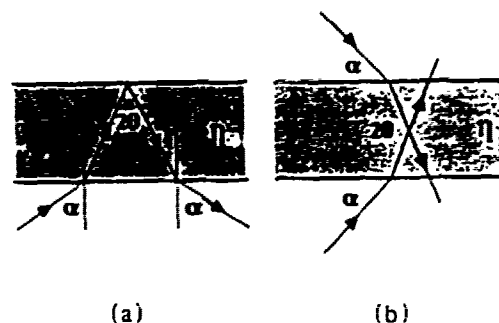
Fig. 1. Writing angle. (a) Transmission hologram. (b) Reflection hologram.

length selectivities for transmission and reflection hologram can be shown as [5]

$$\left(\frac{\delta\lambda}{\lambda}\right)_t = \frac{(\eta^2 - \sin^2\alpha)^{1/2}}{\sin^2\alpha} \frac{\lambda}{d}. \tag{1}$$

and

$$\left(\frac{\delta\lambda}{\lambda}\right)_r = \frac{1}{(\eta^2 - \cos^2\alpha)^{1/2}}. \tag{2}$$

where α is the incident angle and η is the refractive index of the hologram. The normalized wavelength selectivities as a function of incident angle are plotted in fig. 2, where we notice that the wavelength se-
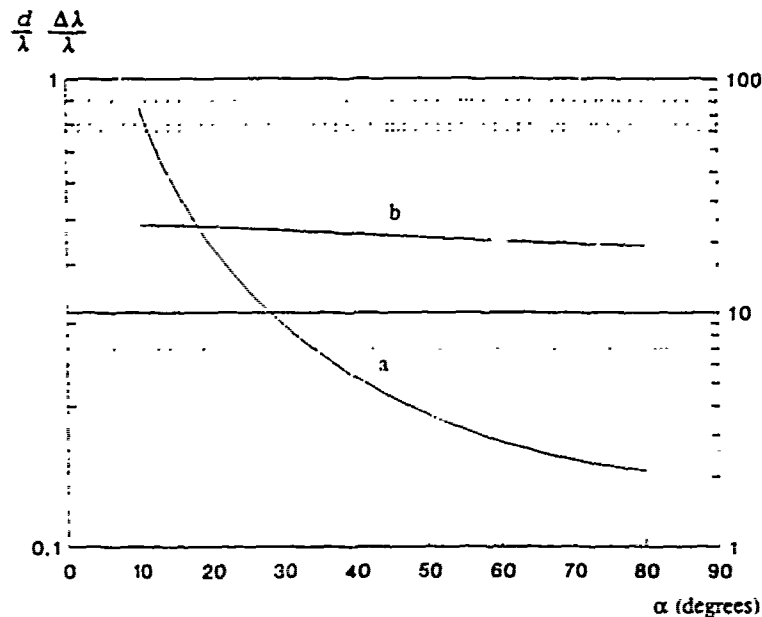
141

$$\frac{d}{\lambda}\frac{\Delta\lambda}{\lambda}$$

Fig. 2. Wavelength selectivity normalized by $(\lambda/d)\eta=2.28$. (a) Transmission hologram (the scale on right). (b) Reflection hologram (the scale on left).

lectivity for reflection hologram is one to two orders of magnitude higher than the transmission hologram.

For a true color hologram construction. a red. a green. and a blue wavelength (e.g., 630 nm. 510 nm. and 440 nm) should be used. in which the relative wavelength spacing is about 1/5. Thus the wavelength selectivity of a color hologram should be smaller than the relative wavelength spacing of the light source. For instance. using 30° incident angle for a transmission hologram. the thickness of the emulsion should be greater than 50λ which is on the order of a few hundred microns. This thickness can easily be obtained by using photorefractive crystal.

To record a color transmission hologram. we used a 1 mm thick LiNbO₃ photorefractive crystal with a writing angle of about 15°. To reduce the noise disturbances caused by the defects of the crystal. a piece of ground glass be placed at the front of a color object transparency. The recording light source is a Lexel Kr-Ar "white-light" laser. which has nine spectral lines distributed within the visible region. The strongest spectral lines of this laser are being 488 nm. 514 nm. and 647 nm wavelengths. Since the spectral lines os late at different transversal modes. spatial filters are used in both the object and the reference

beams to insure the color uniformity. The object-to-reference beam ratio in our experiment is about 1:8, for which a very high quality color hologram is constructed. Fig. 3 shows a color hologram image reconstructed using the "white-light" laser. notice that both vertical and horizontal parallaxes are preserved. Furthermore. by viewing the hologram image at different angles. it does not cause any change in the color. as in contrast with the rainbow holographic imaging.

We stress that. excellent color fidelity can always be obtained by properly controlling the exposure and spectral lines polarizations. Nevertheless. if the exposure is incorrect. color saturation dominated by the strongest spectral line would occur in the hologram image.

Multiplexing construction is a means of high density holographic storage. for which the thickness of photorefractive crystal makes it possible. Fig. 4 shows an optical setup for the multiplexing color hologram construction process using a LiNbO₃. A condensing lens L_1 is used to focus the object beam onto the crystal. The direction of the object beam can be adjusted by positioning L_1. by which various object transparencies can be multiplexing into the photo-

349

Fig. 3. Reconstructed color hologram image.

refractive crystal. In the hologram image reconstruction process, a pinhole is placed at P_1 for the selection of the hologram images, which would be imaged onto plane P_2 by L_2. In our demonstration, we have multiplexed four color holograms in the $LiNbO_3$ crystal. We note that, multi-exposure technique [3] can be used to achieve the same diffraction efficiency for each color hologram image. Furthermore, the size of the pinhole used at P_1 is about $m\lambda F/a$, where $m \times m$ is the number of pixel of the object, a is the size of the object, and F is the focal length of
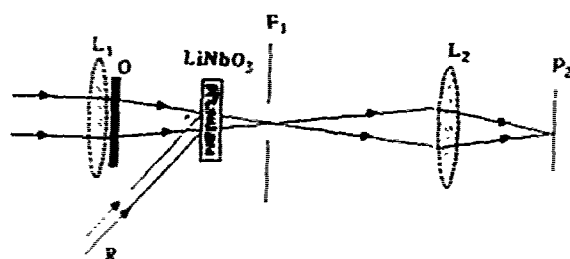
Fig. 4. Geometry for multiplexing color holographic storage.

L_1. It is therefore apparent that if an $N \times N$ pinhole array is used, then N^2 objects can be multiplexed in the photorefractive crystal. Since the pinhole array can be realized by using a programmable spatial light modulator, it is clear that the hologram images can be randomly addressed during the reconstruction process.

In the color hologram recording process, we have found that color crosstalk occurs, i.e., the grating formed at one wavelength is read out by another wavelength. The wavelength selectivities of eqs. (1) and (2) are essentially the width of the main lobe of the frequency detuning response. However, for the dynamic two-wave mixing, the weak sidelobes would have caused serious color crosstalk. Consequently, the color crosstalks are cascaded which will deteriorate the reconstructed image quality. The color crosstalk can be interpreted as the spatial beating phenomena. In fact, we have observed the beating frequencies up to the fifth-order.

Fig. 5 shows a wavevector diagram to represent the beating process. The writing vectors OA and OC are produced by wavelength λ_1, while OB and OD
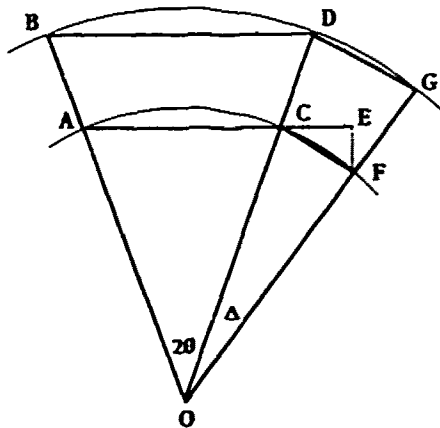
Fig. 5. Wavevector diagram to represent color crosstalk.

are resulted from λ_2. They form grating vectors AC and BC respectively. If the read-out vector is paralleled with respect to vectors OA or OB, then the grating vector BD can only be read out by λ_2. However, if the wavelength selectivity of the hologram is not high enough, or due to the response from the sidelobes, then the grating vector BD can also be read-out by λ_1. In fact, the read-out vector due to crosstalk can be interpreted as the conservation of (wave vector) momentum along the direction of the grating vector. Thus by translating vector BD to AE, and drawing EF, parallel to the bisector of angle (2θ), and intersecting the circle at F, a new wavevector OF can be drawn. We have also seen that the wavevectors OC and OF form a beating grating CF, which can be read out by vector OD of λ_2. Similarly, a new wavevector OG is produced by λ_2 at about the same direction of OF. Since vectors OF and OG would interact with OA and OB, it would generate a secondary set of beatings. The angle of the beating grating

can be computed as follows,

$$\Delta = 2 \tan \theta (\lambda_1/\lambda_2 - 1) . \tag{3}$$

We stress that, for small θ, the angles of beatings would be about the same in different orders.

Fig. 6 shows the cascading beating phenomenon generated by multi-color two-beam mixing in a LiNbO$_3$ photorefractive crystal of 1 mm thick, where the interference angle 2θ is set about $15°$. The brightest spot shown in this figure represents the normal reconstructed spot, while the other spots are generated from the beating phenomenon due to color crosstalks. The color crosstalks are primarily caused by the blue light gratings read-out by green wavelength or vice versa. For example, the angle of the beating grating, Δ, of the right-side spots is measured to be $0.0103 \pm 0.001°$, which is in agreement with the calculated result of $0.0102°$, as obtained from eq. (3). Similarly, the spots on the left-side are generated by the green beating gratings read-out by blue light.

We have demonstrated color holographic imaging using a thick crystal LiNbO$_3$ photorefractive crystal with a "white-light" laser. We have experimentally shown that high quality color hologram image can be obtained with a thick crystal without sacrificing the parallax. Multiplexing color holograms in a thick photorefractive crystal is also discussed, for which we have proposed using a programmable spatial light modulator for the random address. Color crosstalk caused by the grating beating phenomena is also studied. To minimize the color crosstalk, a higher wavelength selectivity (i.e., thick crystal) photorefractive media than normally required should be used.

Fig. 6. Grating beating phenomenon.

References

[1] Y.N. Denisyuk. Soviet Phys. Doklady 7 (1962) 543.
[2] S.A. Benton, J. Opt. Soc. Am. 59 (1969) 1545.
[3] F.T.S. Yu, A. Tai and H. Chen. Opt. Eng. 19 (1980) 666.
[4] H. Kogelnik. B.S.T.J. 48 (1969) 2902.
[5] S. Wu. Q. Song, A. Mayers, D. Gregory and F.T.S. Yu. Appl. Optics 29 (1990) 1118.

APPENDIX 10.20

Compact LCTV Neural Network

a reprint from Applied Optics

Compact optical neural network using cascaded liquid crystal television

Xiangyang Yang, Taiwei Lu. and Francis T. S. Yu

Pennsylvania State University, Electrical Engineering Department, University Park, Pennsylvania 16802.
Received 5 June 1990.

An optical neural network using two tightly cascaded liquid crystal televisions is presented. This new optical architecture offers compactness in size, ease of alignment, higher light efficiency, better image quality, and low cost. The implementation of the autoassociative and heteroassociative memories is given.

Much work has been done to bring about the optical implementation of neural network,[1-7] primarily on the parallelism and massive interconnectivity of optics. The first 2-D optical neural network (2-D ONN) to simulate the Hopfield model was proposed by Farhat and Psaltis in 1986.[2] An LED array was used as the input feature composer, which is optically interconnected to a weighting mask, called the interconnection weight matrix (IWM) or associative memory matrix by a lenslet array. To provide the neural network with learning capability, a fine resolution and high dynamic range programmable spatial light modulator (SLM) are required for the generation of the IWM.

Recently we proposed a different ONN architecture using a high resolution video monitor,[5,6] by which the interconnective part in the Hopfield equation can be optically added. More recently, we constructed an ONN using liquid crystal
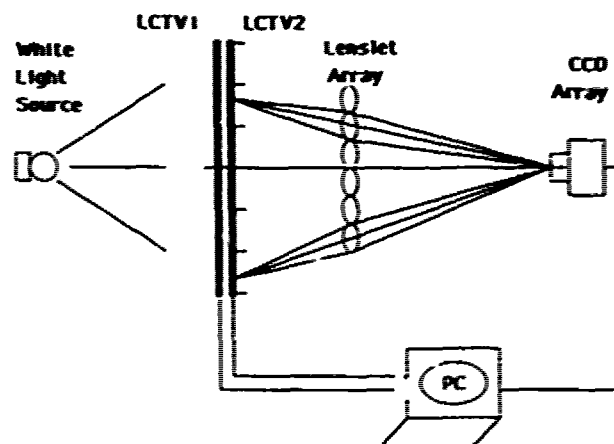
Fig. 1. Compact ONN architecture using cascaded LCTVs.

televisions (LCTVs).[7] In view of the preceding ONN architectures, a large numerical aperture imaging lens, of the order of 1/0.7, is needed to image the lenslet array onto the output charge-coupled device (CCD) detector. The architectures suffer from low light efficiency, high aberration, and larger size due to the imaging lens.

To alleviate these problems, we propose a new compact architecture using tightly cascaded LCTV's, as shown in Fig. 1. Let us recall the iterative operation of a 2-D $N \times N$ neuron network. It can be described by a matrix-vector product operation of Hopfield as given by[2]

$$V_{ik}(n + 1) = f\left[\sum_{i=1}^{N} \sum_{j=1}^{N} T_{ik,ij} U_{ij}(n) \right], \quad (1)$$
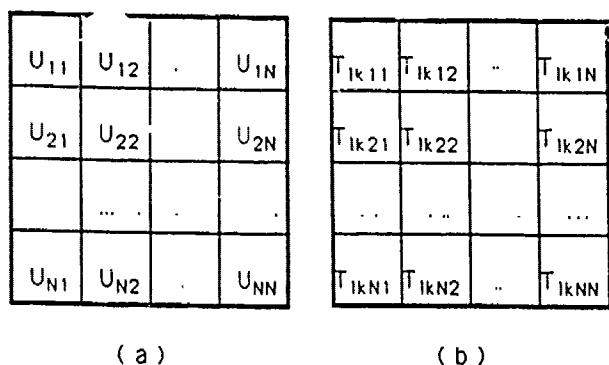
Fig. 2. Display format of a 2-D input pattern and 4-D IWM: (a) input pattern; (b) IWM.

(a)

(a)

(b)

(c)

(b)
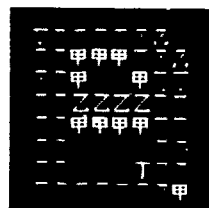

(c)

(d)

(e)
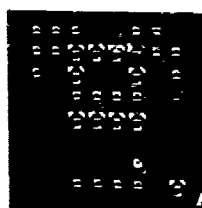
(d)

(e)

Fig. 3 Experimental result for an autoassociative memory. (a) four capital letters stored in the IWM. (b) positive IWM: (c) negative IWM, (d) partial input pattern; (e) reconstructed pattern.

Fig. 4. Experimental result for heteroassociative memory: (a) input–output training set; (b) positive IWM; (c) negative IWM; (d) partial input pattern; (e) output pattern.

where V_{lk} and U_{ij} represent the state of the lkth and ijth neuron in the $N \times N$ neuron space, T_{lkij} is a 4-D IWM, and n represents the nth iteration. Notice that the matrix T can be partitioned into an array of 2-D submatrices T_{lk11}, T_{lk12}, ..., and T_{lkNN}, in which each submatrix is of the $N \times N$ size. Thus the IWM can be displayed as an $N^2 \times N^2$ array representation.

By referring to the proposed architecture of Fig. 1, the input pattern and the IWM are displayed on the cascaded LCTVs, as illustrated in Fig. 2. The light emerging through the cascaded LCTVs is apparently proportional to the product of $T_{lkij}U_{ij}(n)$ in Eq. (1). It can be seen that each subma-

trix $T_{lkij}U_{ij}(n)$ is superimposingly imaged onto the CCD detector by the lenslet array. Thus the output intensity array from the CCD detector is the summation of the submatrices $T_{lkij}U_{ij}(n)$ over i and j. We stress that the summation is essentially carried out optically by the lenslet array. Needless to say that, by thresholding the signals, the result can be fed back to the input LCTV via a microcomputer for the next iteration. Therefore, a closed loop operation can be obtained with the proposed system. Furthermore, with the elimination of the imaging lens, the system can be built in compact size.

Moreover, by cascading the input pattern with the IWM, the output pattern pixels from the ONN remain of the same shape as the input pattern instead of circular shape as in the previous architectures.

An 8×8 neuron neural network using a cascaded Hitachi 12.7-cm (5-in.) color LCTVs is built in the Electro-Optics Laboratory at the Pennsylvania State University. The resolution of the LCTV is 240×480 pixels with a contrast ratio of above 15:1 under white light illum: ion.

In the experimental demonstrations. we have used the interpattern association (IPA) model for the construction of the IWM.[8] Both autoassociative and heteroassociative IWMs have been used in the ONN. An autoassociative

IWM for the English letters A, B, C, and D is shown in Figs. 3(b) and (c). If a partial pattern of A, as illustrated in Fig. 3(d), is presented at the input of the ONN, a recovery of pattern A is obtained after one iteration in our experiment, as shown in Fig. 3(e).

For the heteroassociative memory, four English letters and four Chinese characters are used as the input–output training set, as shown in Fig. 4(a). The heteroassociative IWM is depicted in Figs. 4(b) and (c). Again, if a partial a is presented at the input of the ONN, the corresponding Chinese character is obtained after one iteration, as shown in Fig. 4(e).

In conclusion, we built a compact ONN using tightly cascaded LCTVs. This optical architecture offers the advantages of compact size, easy alignment, higher light efficiency, better image quality, and low cost. Demonstrations of the autoassociative memory and heteroassociative translation are provided.

References

1. N. H. Farhat, D. Psaltis, A. Prata, and E. G. Paek, "Optical Implementation of the Hopfield Model," Appl. Opt. 24, 1469–1475 (1985).
2. N. H. Farhat and D. Psaltis, "Optical Implementation of Associative Memory Based on Models of Neural Networks," in *Optical Signal Processing*, J. L. Horner, Ed. (Academic, New York, 1987), pp. 129–162.
3. B. Macukow and H. H. Arsenault, "Optical Associative Memory Model Based on Neural Networks Having Variable Interconnection Weights," Appl. Opt. 26, 924–928 (1987).
4. J.-S. Jang, S.-W. Jung, S.-Y. Lee and S.-Y. Shin, "Optical Implementation of the Hopfield Model for Two-Dimensional Associative Memory," Opt. Lett. 13, 248–250, (1988).
5. S. Wu, T. Lu, X. Xu, and F. T. S. Yu, "An Adaptive Optical Neural Network Using a High Resolution Video Monitor," Microwave Opt. Technol. Lett. 2, 252–257 (1989).
6. T. Lu, S. Wu, X. Xu, and F. T. S. Yu, "Two-Dimensional Programmable Optical Neural Network," Appl. Opt. 28, 4908–4913 (1989).
7. F. T. S. Yu, T. Lu, and X. Yang, "Optical Neural Network Using Pocket Size Liquid Crystal Televisions," Opt. Lett. 15, 863–865 (1990).
8. T. Lu, X. Xu, S. Wu, and F. T. S. Yu, "Neural Network Model Using Interpattern Association," Appl. Opt. 29, 284–288 (1990).

APPENDIX 10.21

Wavelength Multiplexed Matched Filters

Wavelength multiplexed reflection matched spatial filters using LiNbO3

Francis T.S. Yu, Shudong Wu, Andrew W. Mayers and Sumati Rajan

Pennsylvania State University, Department of Electrical Engineering, 121 Electrical Engineering East, University Park, PA 16802, USA

The wavelength multiplexing properties of a reflection filter are investigated. A photorefractive LiNbO3 crystal was used to construct an experimental filter. The wavelength selectivity and shift invariant properties of the filter were analyzed and experimentally verified.

Recently, multiplexing of filters in photorefractive materials has attracted much attention. By measuring the maximum phase shift achieved in a deep volume hologram, and calculating an exposure schedule for recording multiple equally diffracting holograms, Hong et al. have estimated that over 1000 holograms could be recorded in a 4 mm thick BaTiO3 crystal [1]. The holograms can be recorded in the crystal using angular or wavelength multiplexing techniques. By using a fixed image position, and changing the angular position of the reference beam, Mok et al. have reported recording 500 high resolution, uniformly diffracting volume holograms in a single LiNbO3 crystal [2]. By writing holograms sequentially, and changing the writing wavelength between exposures, multiple holograms can be recorded in the crystal. This wavelength multiplexing technique has the advantage that neither the object nor the reference beam needs to be changed in the writing or reading processes, even though the two processes can be used in conjunction to further increase the multiplexing capabilities. This technique has been used in holographic film in which matched filters were multiplexed [3]. However, the thin film emulsion leads to low wavelength selectivity, and the filters are difficult to produce. In addition, the system was not capable of quick updating of the reference patterns.

In this paper we will investigate wavelength multiplexed matched spatial filters using a LiNbO3 crystal. Referring to the coupled-mode analysis of Kogelnik [4], we will analyze the wavelength multiplexing capacity of reflection type matched spatial filters. By integrating over the thickness of the crystal, as was suggested by Gheen and Cheng [5], we will show that thick crystal filters exhibit high wavelength selectivity as well as retaining a good degree of the shift invariance. Thus, by recording single wavelength filters, and reading-out the filters with a multi-wavelength source, parallel pattern recognition can be performed.

It is well known that volume holograms offer high diffraction efficiency and high angular and wavelength selectivities. Reflection volume holograms have a much higher wavelength selectivity than transmission volume holograms, while the angular selectivity of the holograms remain similar. For this reason, we chose to use reflection volume holograms for the matched filter synthesis. Using the weak coupling approximation, where multiple diffractions are ignored, the wavelength selectivity of the reflection hologram can be expressed as [6]

$$\left(\frac{\Delta\lambda}{\lambda}\right)_r = \frac{1}{\sqrt{n^2-\cos^2\alpha}}\frac{\lambda}{D}, \qquad (1)$$

where λ is the central wavelength of the writing beam, D is the thickness of the crystal, n is the refractive index of the crystal, and α is the external angle shown in fig. 1.
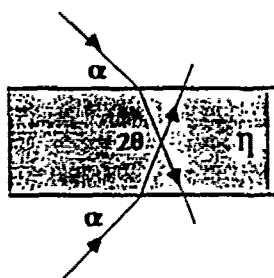
Fig. 1. Writing angle for reflection hologram.

wavevector representations in the crystal can be expressed as.

$$k_0(\lambda_1) = -(n/\lambda_1)\hat{z},$$

$$k_1(\lambda_1) = -\frac{x_1}{\lambda_1}\hat{u} + \frac{n}{\lambda_1}\left(1 - \frac{x_1^2}{2n^2}\right)\hat{z},$$

$$k_2(\lambda_2) = -\frac{x_2}{\lambda_2}\hat{u} + \frac{n}{\lambda_2}\left(1 - \frac{x_2^2}{2n^2}\right)\hat{z},$$

$$k_3(\lambda_2) = \frac{x_3}{\lambda_2}\hat{u} - \frac{n}{\lambda_2}\left(1 - \frac{x_3^2}{2n^2}\right)\hat{z}, \qquad (2)$$

where $\hat{z}$ and $\hat{u}$ are the unit vectors along and normal to the optical axis of the system. respectively. The output correlation signal can be expressed by the following integral equation:

$$R(x_3) = A\left|\int\int\int\int q_1^*(x_1)\, q_2(x_2)\right.$$

$$\left. \times \exp[i2\pi\Delta k\,(u\hat{u} + z\hat{z})]\,dx_1\,dx_2\,du\,dz\right|^2, \qquad (3)$$

where Δk represents the Bragg dephasing vector,

$$\Delta k = k_0 - k_1 + k_2 - k_3. \qquad (4)$$

For simplicity, assume that the size of the crystal is infinite in the $\hat{u}$ direction, therefore the $\hat{u}$ component of the Bragg dephasing wavevector must be zero. From eq. (4), we have,
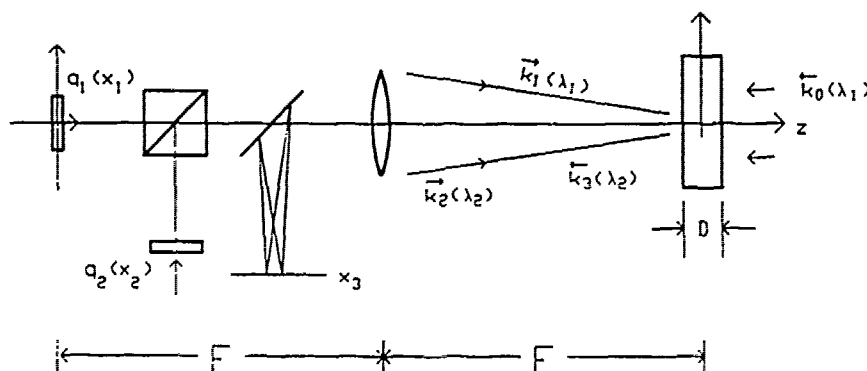
$$x_1/\lambda_1 - (x_2 + x_3)/\lambda_2 = 0. \qquad (5)$$

The shift invariant property of the spatial filter can be determined using layer integration over the thickness of the crystal as proposed by Gheen and Cheng [5]. Fig. 2 shows a schematic diagram for the synthesis of a reflection type matched spatial filter. Without loss of generality, we will use a 1-D representation for the analysis, in which $q_1(x_1)$ represents the reference image that is located at the front focal plane of the transform lens L. For a given point x_1, a wave represented by wavevector $k_1(x_1, \lambda_1)$ is incident at the photorefractive crystal. The reference beam is a plane wave represented by wavevector $k_0(\lambda_1)$, and is incident on the opposite side of the crystal. Thus. it is apparent that a reflection type matched filter can be recorded in the crystal. Let $q_2(x_2)$ be the readout beam representing the input object. which is to be detected. Thus, each point x_2 would produce a readout wavevector $k_2(x_2, \lambda_2)$ for which the reconstructed wavevector is $k_3(x_3, \lambda_2)$, where x_3 denotes the output coordinate plane. The

Fig 2. Geometry for reflection matched filter x_1. x_2. and x_3 are reference. input. and output plane. respectively k_0. k_1. k_2. and k_3 are reference. writing. reading. and readout wavevectors respectively

It follows that

$$x_2 = (1 - \Delta\lambda/\lambda_1)x_1 - x_3 , \qquad (6)$$

where $\Delta\lambda = \lambda_1 - \lambda_2$. Since the relative wavelength difference, $\Delta\lambda/\lambda_1$ is small, the third order terms can be neglected. Thus, the $\hat{z}$ component of the Bragg dephasing vector can be written as

$$\Delta k_z = \frac{1}{n\lambda_1}\left[2n^2 \frac{\Delta\lambda}{\lambda_1} - x_3(x_1 - x_3) \right]. \qquad (7)$$

By substituting eqs. (6) and (7) into eq. (3), and integrating with respect to u, z, and x_2, the correlation signal is

$$R(x_3) = A \left| \int q_1^*(x_1) \, q_2 \left[\left(1 - \frac{\Delta\lambda}{\lambda} \right) x_1 - x_3 \right] \right.$$

$$\left. \times \mathrm{sinc}\left[\frac{\pi D}{n\lambda_1} \left(2n^2 \frac{\Delta\lambda}{\lambda_1} - x_3(x_1 - x_3) \right) \right] \mathrm{d}x_1 \right|^2 . \qquad (8)$$

With reference to the sinc factor, one can evaluate the wavelength selectivity and the shift invariant properties of the thick crystal filter. If the input object is centered with respect to the reference function (i.e., $x_1 = x_2$), then the filter would respond to the wavelength variations that satisfy the inequality of eq. (9),

$$|\Delta\lambda/\lambda_1| < \lambda_1/nD . \qquad (9)$$

This result is essentially the same as was obtained with coupled wave theory of eq. (1) when the beams are perpendicular to the surface of the crystal.

On the other hand, if the read-out and the writing wavelength are identical, then the filter would respond to object shifting that satisfies the inequality of eq. (10),

$$\left| \frac{x_3(x_1 - x_3)}{F^2} \right|_{max} < \frac{\lambda_1}{nD} , \qquad (10)$$

where F represents the focal length of the transform lens.

By substituting eq. (6) into eq. (10), and letting $x_2 = x_1 - \Delta x$, we have

$$|\Delta x(x_1 - \Delta x)|_{max} < \lambda_1 F^2/nD . \qquad (11)$$

This inequality shows that the allowable shift, Δx, is dependent on the halfwidth of the input object $(x_1 - \Delta x)_{max}$, the thickness of the crystal, and the fo-

cal length of the lens. It follows that the wavelength selectivity and shift invariance property are dependent on the thickness of the crystal. In addition, by changing the effective focal length F of the system, a desired shift invariance for a given thickness of the crystal can be obtained [7].

Fig. 3 shows the optical setup that was used for the experiments. The beam splitter separates the beam into a reference beam that enters the crystal from the right, and an object beam that illuminates the input plane, which is transformed by lens L_1 onto a 1 mm thick LiNbO$_3$ crystal. These two beams interfere in the crystal to form a reflection type matched filter. After the filter has been recorded, the reference beam is blocked. and an input object to be correlated with the filter is placed in the input plane. The light scattered by the filter forms an output signal that is transformed by lens L_2 onto the output CCD detector, which is connected to a video monitor for real-time observation.

In the first experimental demonstration, we tested the wavelength multiplexing capabilities of a LiNbO$_3$ crystal. By using $\lambda = 488$ nm, $D = 1$ mm, and $n = 2.25$ in eq. (9), we found that the filter should not respond for wavelength shifts of $\Delta\lambda > 0.106$ nm. We recorded individual matched filters using wavelengths 457.9, 465.8, 488, and 496.5 nm. Reading out these filters with wavelengths different from the recorded wavelength produced no visible output signal. Next, a multiplexed filter using the previously cited wavelengths was recorded of the capital letters D, F, K, and J. Fig. 4a shows the output autocor-
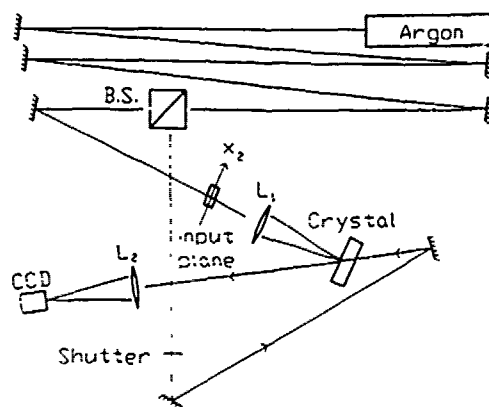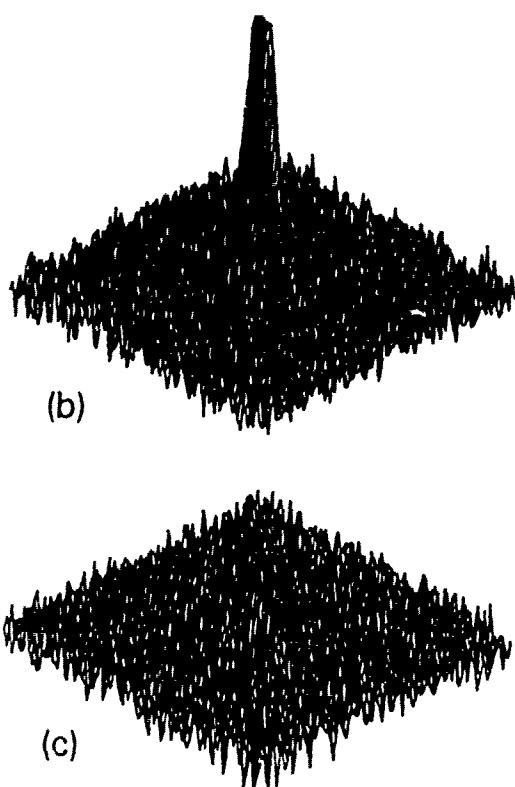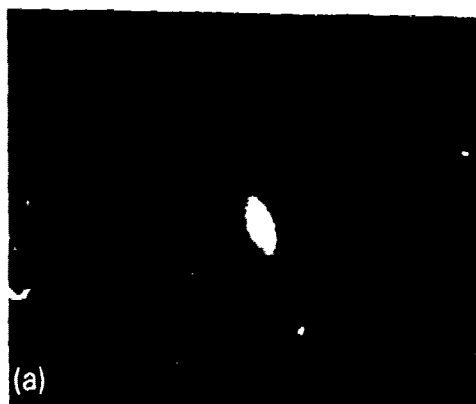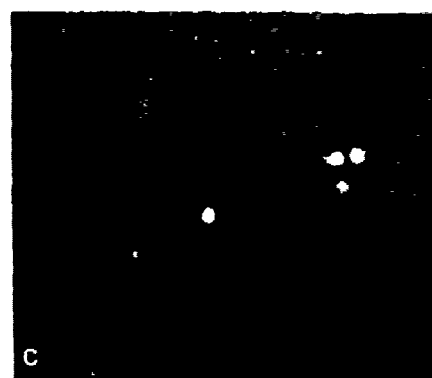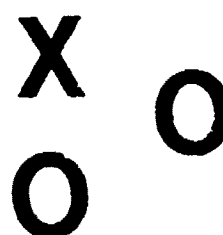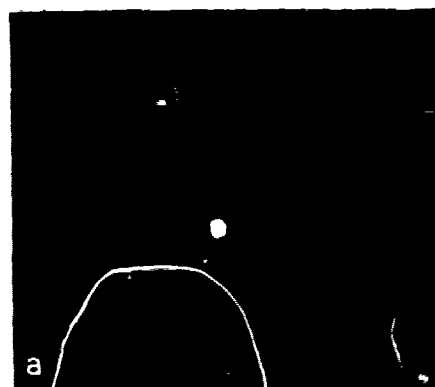

Fig. 3. Experimental setup.

Fig. 5 Shift invariance property of reflection volume filter. (a) Autocorrelation of letter O. (b) Input object consisting of two O's and an X. (c) Output of correlator using input shown in part b.

Fig 4 Output of correlator. (a) Autocorrelation of letter D. (b) Intensity plot of output from part a (c) Crosscorrelation with letter C. using same wavelength.

relation spot of a letter D using wavelength 457.9 nm. The corresponding intensity profile of the autocorrelation signal detected by the CCD camera is shown in fig. 4b. In this figure, we see that a high intensity

correlation peak was obtained. Fig. 4c shows the output cross correlation of this filter with an input object C using a wavelength 457.9 nm. It is apparent that no detectable peak was observed. Thus, we have seen that the wavelength multiplexing capabilities of the crystal filter are in good agreement with our analysis.

In the second experiment the shift invariant property of the crystal filter was verified. By using $\lambda = 488$ nm. $n = 2.25$. $D = 1$ mm. and effective focal length $F = 1$ m. in eq. (10) is clear that $[\Delta x (x_1 - \Delta x)]$ must be less than 216 mm^2. The objects that were used for this experiment have halfwidth of 1.3 mm. for which the shift invariance expression reduces to $\Delta x_{max} < 16$ cm. which is much larger than the beam diameter. A reflection type filter of capital letter O. centered at the input plane was recorded. The corresponding output autocorrelation signal is shown in fig. 5a. Next. a transparency consisting of two O's and an X as shown in fig. 5b was inserted at the input plane of the optical system. The letters were translated about 9 mm with respect to the position of the reference object of the filter. The corresponding output correlation spots are shown in fig. 5c. where two strong correlation spots are observable. as well as a weak cross-correlation distribution. The input beam in this experiment had a usable radius of about ~1.2 cm. and beyond this radius the beam exhibited rapid amplitude variations. Next. the input object was shifted within the beam. and we found that very little amplitude variation in the correlation peak intensity was observed. As before. the experimental results are in good agreement with our analysis.

References

[1] J.H. Hong. P. Yeh. D. Psaltis and D. Brady. Optics Lett. 15 (1990) 344.

[2] F. Mok. M. Tackett and H.M. Stoll. in: Digest of Annual Meeting of the Optical Society of America (Optical Society of America. Washington. D.C.. 1989) p. 76.

[3] S.K. Case. Appl. Optics 18 (1979) 1890.

[4] H. Kogelnik. Bell. Sys. Tech. J. 48 (1969) 2909.

[5] G. Gheen and L.J. Cheng. Appl. Optics 27 (1988) 2756.

[6] S. Wu. Q. Song. A. Mayers. D. Gregory and F.T.S. Yu. Appl. Optics 29 (1990) 1118.

[7] F.T.S. Yu. S. Wu and Sumati Rajan. Compact joint transform correlator using a thick photorefractive crystal. submitted to Optics Lett.

APPENDIX 10.22

Space-Time-Sharing Neural Network

Space–time-sharing optical neural network

Francis T. S. Yu, Xiangyang Yang, and Taiwei Lu

Department of Electrical and Computer Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

A space–time-sharing optical neural network for implementing a large-scale operation is presented. If the interconnection weight matrix is partitioned into an array of submatrices, a large space–bandwidth pattern can be processed with a smaller neural network. We show that the processing time increases as a square function of the space–bandwidth product of the pattern. To illustrate the space–time-sharing operation, experimental and simulated results are provided.

The massive interconnection and parallel processing capabilities of optics offer advantages for optical computing. In recent years, much effort has been applied for the optical implementation of neural networks.[1-10] To provide the neural network with learning ability, spatial light modulators (SLM's), such as a video monitor[7,8] and a liquid-crystal television[9,10] (LCTV), have been used as the programmable devices. In a fully interconnected neural network, every neuron at the input plane is connected to all the neurons at the output plane. For example, 1000 neurons would require a million interconnections. Thus an extremely high-resolution SLM is required for the massive interconnection. However, the resolution of the currently available SLM's is rather limited. This poses an obstacle in developing an optical neural network for large-scale operation. In this Letter we use a space–time sharing technique to alleviate this constraint.

For an $N \times N$ neuron network, the interative equation can be described by[2]

$$V_{lk}(n + 1) = \mathbf{f}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} T_{lkij}U_{ij}(n) \right]. \quad (1)$$

where n denotes the nth iteration, $\mathbf{f}(\cdot)$ represents a nonlinear operator, V_{lk} and U_{ij} represent the state of the lkth and ijth neurons, respectively, and T_{lkij} is the connection strength from the lkth to the ijth neuron. $[T_{lkij}]$ is known as the interconnection weight matrix (IWM) and can be partitioned into an array of $N \times N$ submatrices.[2,8]

An optical neural network (Fig. 1) is used in our discussion, in which the IWM and the input pattern are displayed onto LCTV1 and LCTV2, respectively. Each lens in the lenslet array images a specific submatrix onto the input LCTV2 to establish the proper interconnections. The lenslet array is imaged by the imaging lens onto the output plane. The overall intensity transmitted by each lens in the lenslet array is picked up by the charge-coupled-device (CCD) detector as the output value. We let the resolution LCTV1 be limited by $R \times R$ pixels, and the lenslet array is equal to $L \times L$ neurons. If the size of the IWM is larger than the resolution of LCTV1, i.e., $N^2 > R$, the

IWM cannot be fully displayed onto LCTV1. We discuss in the following cases how a small neural network may accommodate a large-size IWM:

For $N^2 > R$ and $LN < R$, we let $D = \text{int}(N/L)$, where $\text{int}(\cdot)$ is the integer function. The IWM can be partitioned into $D \times D$ sub-IWM's, and each sub-IWM consists of $L \times L$ submatrices of $N \times N$ size, as shown in Fig. 2. To complete the iterative operation of Eq. (1), $D \times D$ sequential operations of the sub-IWM's are required. Thus a smaller neural network can handle a larger space–bandwidth product (SBP) input pattern through sequential operation of the sub-IWM's.

For $N^2 > R$ and $LN > R$, we let $D = \text{int}(N/L)$ and $d = \text{int}(LN/R)$. In this case the submatrices within the sub-IWM's are further divided into $d \times d$ smaller submatrices, and each submatrix is $(N/d) \times (N/d)$ in size, as shown in Fig. 3. The iterative equation can be written as

$$V_{lk}(n + 1)$$
$$= \mathbf{f}\left\{ \sum_{p=0}^{d-1}\sum_{q=0}^{d-1}\left[\sum_{i=pd+1}^{(p+1)(N/d)}\sum_{j=qd+1}^{(q+1)(N/d)} T_{lkij}U_{ij}(n) \right] \right\}. \quad (2)$$

The input pattern is also partitioned into $d \times d$ submatrices, and each submatrix is $(N/d) \times (N/d)$ in size. Thus by sequentially displaying each of the IWM submatrices with respect to the input submatrices onto LCTV1 and LCTV2, respectively, one can process an extremely large SBP pattern with a small neural net-
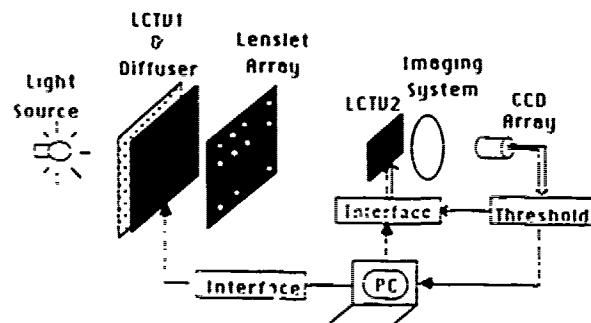

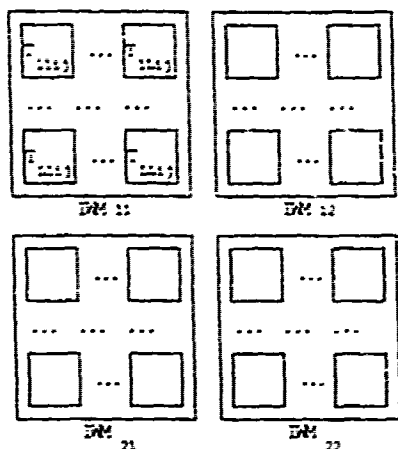Fig. 1. Schematic diagram of the optical neural network.

Fig. 2.   Partition of the IWM into $D \times D$ sub-IWM's ($D = 2$).
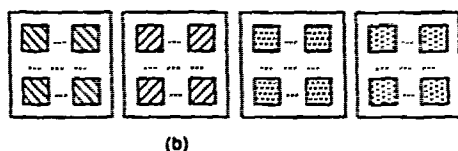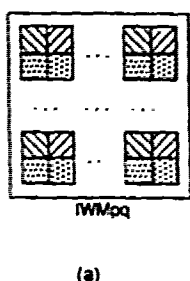

(a)

(b)

Fig. 3.   Partition of a sub-IWM into $d \times d$ submatrices ($d = 2$). (a) The pqth sub-IWM. (b) $d \times d$ smaller submatrices.

work.   It is trivial that the price we paid is prolonging the processing speed by $D^2 \times d^2$ times.

Generally, the processing time increases as the square function of the SBP, of the input pattern, as given by

$$T_2 = [(N_2 \times N_2)/(N_1 \times N_1)]^2 T_1 = (N_2/N_1)^4 T_1,$$

$$(N_2 > N_1 > = L),   (3)$$

where T_2 and T_1 are the processing times for the input patterns with $N_2 \times N_2$ and $N_1 \times N_1$ resolution elements, respectively   For instance, if the resolution elements of the input pattern increase four times in each dimension, i.e., $N_2 = 4N_1$, the processing time would be $4^4 = 256$ times longer.

For experimental demonstrations we show that patterns with 12×12 resolution elements can be processed using the 6×6 neuron network shown in Fig. 1. A 240×480 element Hitachi color LCTV (LCTV1) is used for displaying the IWM.   Since each color pixel is composed of red, green, and blue elements, the resolution is actually reduced to 240×160 pixels.   In the experiment, however, we used 2×2 pixels for each

interconnection weight, with which the resolution of the LCTV1 is essentially reduced to 120×80 elements.

By referring to $N = 12$, $L = 6$, and $R = 80$ and noting that $N^2 = 144 > R$ and $LN = 72 < R$, we have $D = \text{int}(N/L) = 2$.   The IWM can be divided into a 2×2
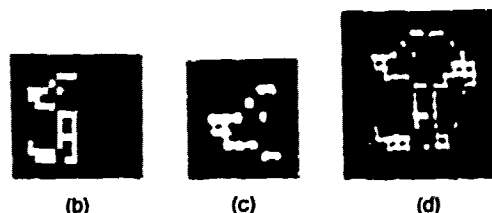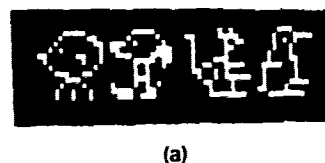

(a)

(b)   (c)   (d)

Fig. 4.   Experimental results of processing 12×12 element patterns on a 6×6 optical neural network. (a) Four reference patterns stored in the IWM, (b) the partial input patterns, (c) one of the four 6×6 suboutput arrays, (d) the composed output pattern.
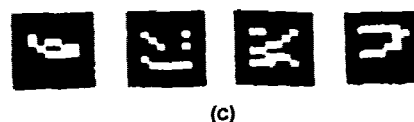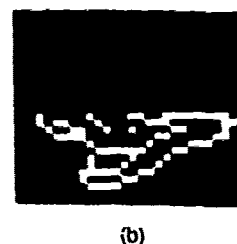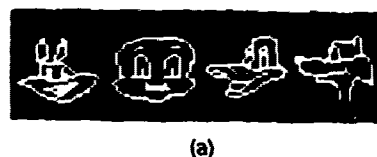

(a)

(b)

(c)

(d)

Fig. 5   Computer-simulated results of processing 24×24 element patterns on a 6×6 optical neural network. (a) Four reference patterns stored in the IWM, (b) the partial input pattern, (c) four of sixteen 6×6 suboutput arrays, (d) the composed output pattern.

158

sub-IWM array and displayed one by one onto LCTV1. If an input pattern is displayed on LCTV2, the signals collected by the CCD camera can be thresholded and then composed to provide an output pattern, which has a SBP four times larger than that of the optical neural network.

One of the experimental results is shown in Fig. 4. The training set is shown in Fig. 4(a), and each of these patterns is limited by a 12×12 pixel matrix. Figure 4(b) shows a part of the second image with 12×5 pixels blocked as the input pattern. The sub-IWM's are then sequentially displayed, one by one, onto LCTV1. Different parts of the output pattern are obtained, and one of the output parts is shown in Fig. 4(c). The final output pattern recalled by the optical neural network is given in Fig. 4(d).

To demonstrate the larger-scale operation further, a 24×24 neuron IWM is used. Since $N = 24, L = 6, R = 80, N^2 = 576 > R$, and $LN = 96 > R$, we take $D \operatorname{int}(N/L) = 4$ and $d = \operatorname{int}(LN/R) = 2$. Thus the IWM is partitioned into 4×4 sub-IWM's, and each sub-IWM is divided into 2×2 smaller submatrices, as illustrated in Fig. 3. In this case the input pattern is also divided into 2×2 submatrices, and each submatrix is 12×12 in size. It is apparent that by sequentially displaying the submatrices of IWM and the input submatrices onto LCTV1 and LCTV2, respectively, a 24×24 output pattern can be obtained.

Figure 5(a) shows four 24×24 pixel images as the training set. A partial pattern of the third image, shown in Fig. 5(b), is used as the input pattern, which is divided into 2×2 matrices of a 12×12 size during the processing. Four of sixteen 6×6 output parts are shown in Fig. 5(c). In going through all the partitions of the sub-IWM's, we compose an output pattern of the image, as shown in Fig. 5(d). The whole process takes $D^2 \times d^2 = 64$ operations of the optical neural network.

In conclusion, we have used a space–time-sharing technique for large-scale neuron operation. We have shown that to achieve a large SBP of the system, additional expenditure of processing time is needed. The amount of processing time increases as the square function of the SBP of the input pattern. For experimental demonstration, we have implemented the processing of 12×12 element patterns on a 6×6 neuron network. The processing of 24×24 element patterns on a 6×6 neuron network has been further demonstrated by computer simulation.

References

1. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, Appl. Opt. 24, 1469 (1985).
2. N. H. Farhat, and D. Psaltis, in *Optical Signal Processing*, J. L. Horner, ed. (Academic, New York, 1987), p. 129.
3. B. Macukow and H. H. Arsenault, Appl. Opt. 26, 924 (1987).
4. N. Farhat, Appl. Opt. 26, 5093 (1987).
5. R. A. Athale and C. W. Stirk, Opt. Eng. 28, 447 (1989).
6. D. Psaltis, D. Brady, X.-G. Gu, and K. Hsu, in *Optical Processing and Computing*, H. H. Arsenault, ed. (Academic, San Diego, Calif., 1989), p. 251.
7. S. Wu, T. Lu, X. Xu, and F. T. S. Yu, Microwave Opt. Technol. Lett. 2, 252 (1989).
8. T. Lu, S. Wu, X. Xu, and F. T. S. Yu, Appl. Opt. 28, 4908 (1989).
9. F. T. S. Yu, T. Lu, X. Yang, and D. A. Gregory, Opt. Lett. 15, 863 (1990).
10. X. Yang, T. Lu, and F. T. S. Yu, Appl. Opt. 29, 5223 (1990).

APPENDIX 10.23

Binary Phase Only JTC

BINARY PHASE ONLY JOINT TRANSFORM CORRELATOR

Francis T. S. Yu and Toshio Nagata
The Pennsylvania State University
Electrical Engineering Department
University Park, Pennsylvania 16802

KEY TERMS

ABSTRACT

A joint transform hybrid optical correlator as applied to binary phase only filtering for pattern recognition is proposed. The interference images of the joint Fourier transformation, due to the reference and the input targets, are captured by a CCD camera feeding into a microcomputer for binarization. A binarized phase only filter, in principle, can be written onto a spatial light modulator (SLM). By simple coherent illumination of

SLM, correlation operation of the reference and the input targets can be obtained. A computer simulation for the target embedded in an additive Gaussian noise is provided.

A programmable joint transform correlator for optical pattern recognition has been discussed in recent papers [1, 2]. In a joint transform correlator, both the reference image and the input target are generated in the input plane of a coherent optical processor. A transform lens will perform the joint transform of the input objects. Due to coherent illumination. uniform carrier interferometric fringes are produced in the joint transform spectrum. The interferometric fringes can be extracted by a square law converter (e.g., film, CCD camera, etc.). Then by simple coherent readout of the joint power spectral distribution, the joint correlation of the input object function can be obtained through a second Fourier transform lens.

In principle. a joint transform correlator is capable of performing any shift-invariant linear operations by using the impulse response of the filter function which is displayed as a reference image [3]. However, for the convolution operation, the reference image must be the complex conjugate of the inverted impulse response function. With the current stage of technology, the spatial light modulator (SLM) is applicable only t real impulse responses due to the lack of a complex image display device. Furthermore, low dynamic range and slow response of the square law converters give rise to another practical restriction. One possible realization of the joint transform correlator is to employ an electrooptic device as a square law converter (e.g., CCD camera). The detected joint transform power spectrum is then displayed on a binary SLM for the second transformation.

In this paper, we propose a joint transform hybrid optical correlator using a binary phase only filter [4, 5]. The system takes advantage of the controllability of the electrooptic devices and the high speed optical processing operation. As is shown in Figure 1, the proposed joint transform correlator involves gray level image displaying devices and a binary phase spatial light modulator both of which can be addressed by a microcomputer. The image displaying devices can be

implemented by using liquid crystal televisions (LCTV) [6]. The phase only filter can be synthesized in a programmable spatial light modulator, e.g., a magnetooptic spatial light modulator (MOSLM) [5]. We note that the joint transform interference fringes can be binarized by a microcomputer. The basic objective of binarizing the fringes is to synthesize a binary phase only joint power spectrum that can be recorded at the MOSLM.

In Figure 1, we assume the reference image $f_1(x, y)$ and the input target $f_2(x, y)$ are written on the LCTV. Under coherent illumination, the complex light distribution $\therefore$ the Fourier plane P_1 is given by

$$G(p, q) = F_1(p, q)e^{jaq} + F_2(p, q)e^{-jaq} \qquad (1)$$

$$F_1(p, q) = \mathscr{F}[f_1(x, y)] \qquad (2)$$

$$F_2(p, q) = \mathscr{F}[f_2(x, y)] \qquad (3)$$

where $\mathscr{F}$ denotes the Fourier transform and $2a$ is the center-to-center separation of the input objects $f_1(x, y)$ and $f_2(x, y)$. We stress that the purpose of the lens L_M is to magnify the joint transform spectrum and that the stop behind the lens L_M selects only a certain bandwidth.

The image captured by the CCD camera can be written as

$$I(p, q) = |G(p, q)|^2$$

$$= |F_1(p, q)|^2 + |F_2(p, q)|^2$$

$$+ F_1(p, q)F_2^*(p, q)e^{j2aq}$$

$$+ F_1^*(p, q)F_2(p, q)e^{-j2aq} \qquad (4)$$

where the superscript asterisk represents the complex conjugate and (p, q) is the spatial frequency coordinate system. For simplicity, the effects of the magnification and the DC stop are neglected in the above equation. It is evident that Equa-
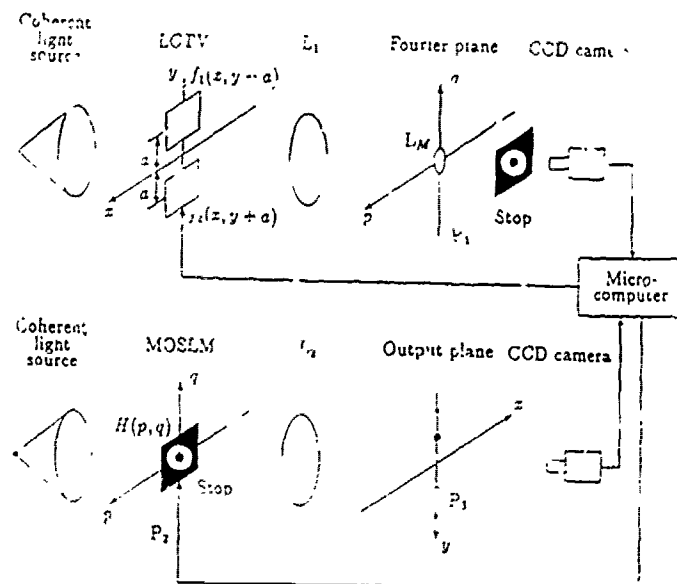

Figure 1 A phase only joint transform programmable optical correlator L_1 and L_2. Fourier transform lenses. L_M magnification lens

tion (4) is the joint power spectral distribution of $F_1(p,q)$ and $F_2(p,q)$. Alternatively, Equation (4) can be written as

$$I(p,q) = |F_1(p,q)|^2 - |F_2(p,q)|^2$$
$$- 2|F_1(p,q)||F_2(p,q)|\cos[2aq + \Delta\phi(p,q)]$$
(5)

where

$$\Delta\phi(p,q) = \phi_1(p,q) - \phi_2(p,q)$$
(6)

and $\phi_1(p,q)$ and $\phi_2(p,q)$ are the phase distribution of $F_1(p,q)$ and $F_2(p,q)$, respectively. It is however interesting to note that if the input objects are identical, i.e., $f_1(x,y) = f_2(x,y)$, then Equation (5) can be reduced to the form

$$I(p,q) = 2|F_1(p,q)|^2(1 + \cos[2 iq])$$
(7)

where the cosine term represents uniform interference fringes of the joint power spectral distribution. On the other hand, if the input objects are different, $\Delta\phi(p,q)$ is not zero and it is a rather complicated function. The cosine factor in Equation (5) no longer provides a strict uniform fringe pattern. Thus it will eventually produce lower correlation peaks through the inverse Fourier transformation.

The fringe pattern captured by the CCD camera can be fed into a microcomputer for binarization with a thresholding procedure. The binarized fringe pattern, in principle, can be written onto a MOSLM to obtain a binarized phase only fringe pattern [5], such that the transmittance function of the MOSLM would be

$$H(p,q) = \begin{cases} e^{+j\pi} & I(p,q) \geq I_T \\ e^{-j\pi} & I(p,q) < I_T \end{cases}$$
(8)

here I_T is the binarizing threshold level. In reality, the interference fringe structure has a Gaussian shaped pedestal component. A stop with a ring window would pass only the uniform pedestal fringe pattern. Then the lowest end of the dynamic range of the A/D converter for the video signal may be selected as the threshold level. Since $e^{\pm j\pi} = \cos(\pm\pi) = \pm 1$, Equation (8) can also be written as

$$H(p,q) = \begin{cases} +1 & I(p,q) \geq I_T \\ -1 & I(p,q) < I_T \end{cases}$$
(9)

If the input objects are identical, then $H(p,q)$ reduces to a one-dimensional bipolar rectangular function with a spatial frequency equal to

$$f = \frac{2a}{f\lambda}$$
(10)

where f is the focal length of the transform lens and λ is the illuminating wavelength. Thus $H(p,q)$ can be expanded into a Fourier cosine series as given by

$$H(p,q) = \sum_{n=1}^{n=-\infty} b_n \cos[2anq]$$
(11)

where b_n are the Fourier coefficients of $H(p,q)$.

If the binary phase distribution of Equation (11) is read out by coherent illumination, as shown in Figure 1, the output complex light distribution can be shown to be

$$g(x,y) = \frac{1}{2} \sum_{n=-\infty}^{n=+\infty} b_n \delta(y - 2an) \qquad b_{-n} = b_n \quad (12)$$

This yields a series of spots. The presence of a first order spot indicates a match. On the contrary, if $f_1(x,y)$ differs from $f_2(x,y)$, the spot would be absent. Thus we see that a binary phase only joint transform correlator can be used for optical pattern recognition.

In the following, we should provide computer simulation tests for the proposed binary joint transform correlator (BJTC). The 128×128 pixel reference and target images are loaded from disk files and written onto the upper and the lower halves of the input plane. We assume that the A/D converter for the video signal has eight bit resolution. To avoid the conversion error, we may discard the least significant bit. The threshold level was chosen to be the lowest end of the dynamic range of the A/D converter such as

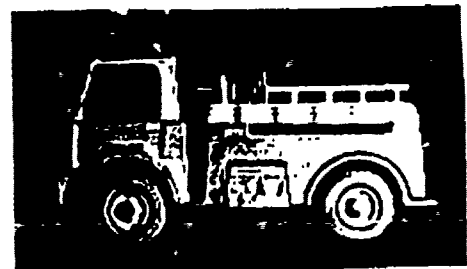$$I_T = \frac{\max_{p,q}[I(p,q)]}{128}$$
(13)

Figure 2(a) shows a reference image. Figures 2(b) and (c) represent target images. Figure 2(b) was obtained by adding an additive Gaussian noise to the reference, with a signal-to-
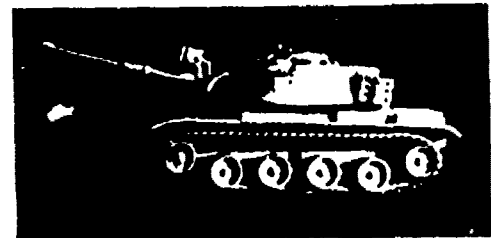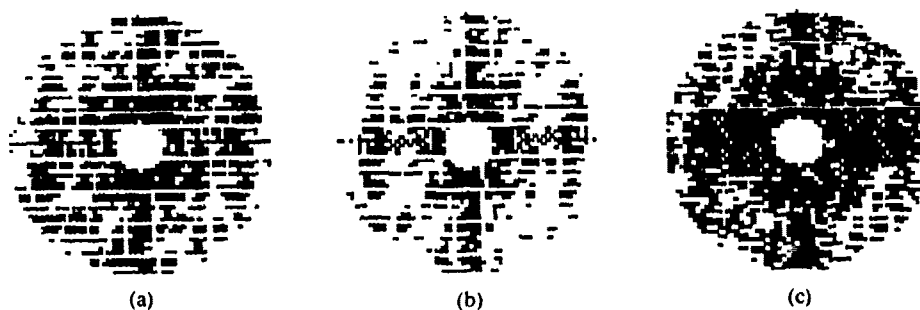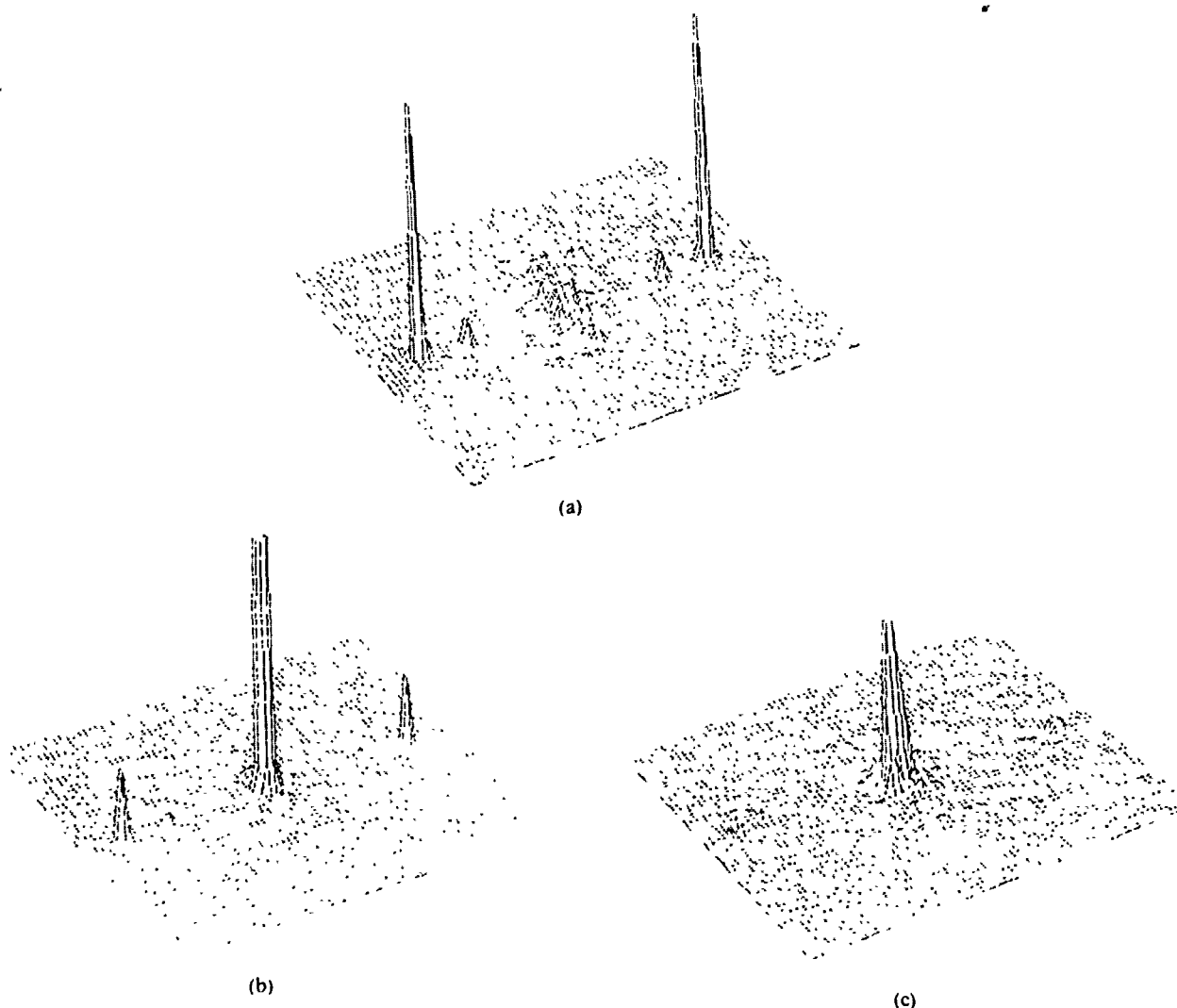

(a)
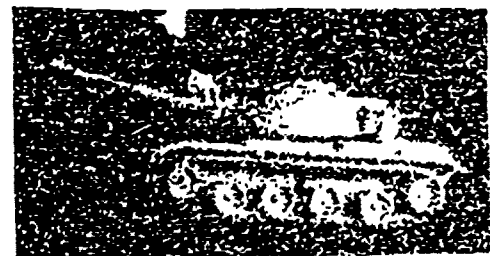
(b)

(c)

Figure 2 Input images: (a) reference image of a tank; (b) target embedded in noise (SNR = 1 dB); (c) target image of a truck

Figure 3 Binarized interference fringes. The DC component is eliminated by a circular stop. (a) Noiseless case of Figure 2(a); (b) noisy case of Figure 2(b); (c) different objects of Figures 2(a) and (c)

(a)

(b)

(c)

Figure 4 Output correlation peaks: (a) noiseless case of Figure 3(a), (b) noisy case of Figure 3(b); (c) different objects of Figures 2(a) and (c)

noise ratio (SNR) equal to 1 dB. Figure 3 shows the computer simulated binarized joint transform power spectrum for noiseless, noisy, and different objects. The SNR for the noisy target is about 1 dB. A uniform spatial frequency fringe pattern can readily be seen in Figure 3(a). However, uniformity of the fringes is somewhat deteriorated under the noisy environment as can be seen in Figure 3(b). No clear fringes can be observed in Figure 3(c). The corresponding output correlations, simulated by digital computer, are shown in Figure 4. A compari-

son between Figures 4(a) and (b) shows that the correlation peaks reduce to about a half for an input SNR = 1 dB. As is shown in Figure 4, no correlation peak was obtained for the cross correlation of different images.

In conclusion, we have investigated a binary phase only joint transform optical correlator for optical pattern recognition. The major advantage of the proposed system must be the avoidance of synthesizing a spatial matched filter. The combination of CCD camera and MOSLM may have an advantage

over the optically addressed SLM such as flexibility of threshold level selection. We have shown that high correlation peaks can be readily obtained under the noisy environment.

REFERENCES

1. F T. S. Yu and X. J. Lu. "A Real-Time Programmable Joint Transform Correlator." *Opt. Commun.*. Vol 52. 1984. p. 47.
2. F. T. S. Yu and J. E. Ludman. "Microcomputer-Based Programmable Optical Correlator for Automatic Pattern Recognition and Identification." *Opt. Lett.*. Vol. 11. 1986. p 395.
3. F. T. S. Yu and J. E. Ludman. "Joint Fourier Transform Processor." to be published.
4. J. L. Horner and P. D. Gianino. "Phase-Only Matched Filtering." *Appl. Opt.*. Vol. 23. 1984. p. 812.
5. D. Psaltis. E. G. Peak. and S. S. Venkatesh. "Optical Image Correlation with a Binary Spatial Light Modulator." *Opt. Eng.*. Vol. 23. 1984. p. 698.
6. F. T. S. Yu. T. W Lin. and D. A. Gregory. "Adaptive Real-Time Pattern Recognition Using a Liquid Crystal TV Based Joint Transform Correlator." *Appl. Opt.*. Vol. 26. 1987. p 1370.